



HAL
open science

**Experimental evidence for short term directional
selection of epigenetic trait variation Epigenetic
response to selection Authors**

Benoit Pujol, Mathieu Latutrie, Nelia Luviano, Pierick Mouginot, Jesaelle
Piquet, Sara Marin, Stéphane Maury

► **To cite this version:**

Benoit Pujol, Mathieu Latutrie, Nelia Luviano, Pierick Mouginot, Jesaelle Piquet, et al.. Experimental evidence for short term directional selection of epigenetic trait variation Epigenetic response to selection Authors. 2024. hal-04777459

HAL Id: hal-04777459

<https://hal.science/hal-04777459v1>

Preprint submitted on 12 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2 **Title**

- 3 • Experimental evidence for short term directional selection of epigenetic trait variation
4 • Epigenetic response to selection

5
6 **Authors**

7 Benoit Pujol,^{1*} Mathieu Latutrie,¹ Nelia Luviano-Aparicio,¹ Pierick Mouginot,¹ Jésaëlle
8 Piquet,¹ Sara Marin,¹ Stéphane Maury²

9
10 **Affiliations**

11 ¹PSL Université Paris: EPHE-UPVD-CNRS, UAR 3278 CRIOBE, Université de
12 Perpignan, 52 Avenue Paul Alduy, CEDEX 9, 66860 Perpignan, France.

13 ²INRAE, EA1207 USC1328 Laboratoire de Biologie des Ligneux et des Grandes
14 Cultures, Université d'Orléans, F-45067 Orléans, France.

15 *Correspondence to: benoit.pujol@univ-perp.fr

16
17 **Abstract**

18 Evolution by natural selection can occur when organisms harbor genetically inherited
19 phenotypic variation, and phenotypic variants have differential fitness. Stable
20 transgenerational epigenetic variation also exists for fitness-related traits and theory
21 predicts that selection can act on this variation alone without a contribution of genetic
22 variation. Here, we artificially selected for divergent rosette size and plant height in
23 experimental *Arabidopsis thaliana* populations harboring DNA methylation
24 polymorphism in an identical genetic background. We found significant epi-allele
25 frequency changes in response to selection. Our results show how selection changed
26 population trait values and their epigenetic basis in real time, over one generation of
27 selection. Our results imply the role of the transgenerational epigenetic variation of
28 populations as an additional source of short-term adaptive potential.

29
30 **Teaser**

31 Trait based selection experiment in *Arabidopsis* recombinant inbred lines drives
32 epigenetic changes in response to selection.

33
34 **MAIN TEXT**

35
36 **The manuscript should be a maximum of 15,000 words.**

37
38 **Introduction**

39 Evolution by natural selection occurs when organisms express heritable variation for
40 fitness. As a result, fitness-related traits change in response to selection (Walsh and Lynch
41 2018). The heritable trait variation upon which selection can act is typically equated to
42 variation in the DNA sequence of bases in most studies of adaptive evolution. However,
43 transgenerational epigenetic variation exists for fitness-related traits (Kronholm et al.
44 2017, Muyle et al. 2021, Stajic et al. 2019). Under the assumption that epigenetic variants
45 are stable across generations, theory predicts that selection can act on this heritable
46 variation alone, without any contribution of genetic variation (Bonduriansky and Day
47 2009, Danchin et al. 2011, Jablonka et al. 2014).

48 Epigenome-wide association studies (EWAS) have been successfully used in the past
49 (Ong-Abdullah et al. 2015) to identify population divergence based on heritable DNA

50 methylation patterns linked to phenotypic traits. The presence of epigenetic divergence
51 between populations that have a variable genomic background is difficult to interpret as
52 strong evidence that selection can shape standing epigenetic variation independently from
53 DNA sequence variation but sheds light on this possible scenario.

54 The stakes are high for evaluating the potential role of epigenetic variation as a source of
55 variation upon which selection can act. This is because the epigenetic response to
56 selection may participate to adaptation of wild populations confronted to environmental
57 constraints. Selection of standing epigenetic variation could also participate to breeding
58 strategies (Kakoulidou et al. 2021). Epigenetic breeding might be rather straight forward
59 in vegetatively propagated crops (Latutrie et al. 2019), and in seed crops if the response to
60 selection can be predicted from the epigenetic heritability and the selection differential, as
61 usually done on the basis of the standing genetic variation of traits (Walsh and Lynch
62 2018). Another stake lies in the debate around the evolutionary significance of epigenetic
63 variation (Charlesworth et al. 20017, Quadrana and Colot 2016, Richards et al. 2017,
64 Verhoeven et al. 2016). This debate roots in the confounded genetic and epigenetic
65 variations of traits that are technically and statistically difficult to separate (Thomson et al.
66 2018). We therefore tested for an epigenetic signature of selection by using a selection
67 experiment in artificial plant populations displaying DNA methylation polymorphism in
68 the same genetic background.

69 Here, we tested the hypothesis that directional selection can act on heritable epigenetic
70 trait variation in an identical genetic background by conducting artificial selection
71 experiments in *Arabidopsis thaliana*. There is growing evidence for epigenetic variation,
72 notably DNA methylation, at quantitative traits in animals and plants (Cortijo et al. 2014,
73 Noshay et al. 2021, Vogt 2021). This is notably the case of size traits and flowering time
74 in *A. thaliana* (Cortijo et al. 2014). Size traits and flowering time strongly influence the
75 fitness of most plants in nature (Halbritter et al. 2018). There is evidence for the heritable
76 epigenetic variation of these traits associated with differentially methylated regions
77 (DMRs) that are stably inherited across generations in epiRILs (Johannes et al. 2009,
78 Reinders et al. 2009, Roux et al. 2011, Zhang et al. 2018). EpiRILs are a set of 123 fixed
79 homozygous lines, descending from an F2 population, that were obtained by an initial
80 cross of isogenic parents with different DNA methylation profiles (Col-0 x *Col-ddm1-2*)
81 (Johannes et al. 2009). The mutation in the DDM1 gene (*ddm1*) – a chromatin remodeler
82 involved in DNA methylation – induced hypomethylation of 126 clusters of cytosines
83 (DMRs). We therefore expect that a selection differential imposed to parental plants on
84 the basis of their size and flowering time which variation would be caused by DMRs, e.g.,
85 by allowing to breed only bigger or late flowering epiRILs, would produce the expected
86 directional trait changes in the progeny.

87 Novel phenotypes can be selected in *A. thaliana* populations that only differ with respect
88 to epigenetic variation in a fixed genomic background (Schmid et al. 2018). These
89 changes were stably inherited over at least two generations in the absence of selection and
90 correlated with changes in DNA methylation and transcription. Potential genetic
91 differences that could have occurred over the 7-8 generations of this experiment were
92 excluded by resequencing the complete genome of several individuals. The epigenetic
93 divergence found between lab populations exposed to fragmented and disturbed landscape
94 environments and natural plant populations could therefore be attributed to selection
95 (Schmid et al. 2018). Today, the time is ripe to reach a better understanding of the
96 epigenetic response to selection. Whether directional selection shapes the standing

97 epigenetic variation of traits as it would do with the standing genetic variation of traits
98 remains to be clarified. Confirming this hypothesis would confirm that the genetical
99 theory of natural selection also applies to epigenetic variation and unlock new potential for
100 breeding strategies.

101 First, building upon the available knowledge, we described DNA methylation
102 polymorphism acknowledged to be stable across generations (Cortijo et al. 2014, Johannes
103 et al. 2009) and phenotypic variation in two different artificial *A. thaliana* epiRIL
104 populations. We grew these two populations made of two totally different sets of *A.*
105 *thaliana* epiRILs in similar conditions and at different times, in order to repeat the
106 experiment in two different populations. Second, independently in each population and for
107 each trait, we artificially generated a fitness differential between plants by selecting them
108 based on their morphological and reproductive trait values. We selected in two different
109 directions - for upper and lower trait values – and applied two different selection
110 intensities - weak and strong - to explore the potential properties and limits of epigenetic
111 selection. Third, we estimated the phenotypic response to selection in the next generation
112 of plants and assessed whether changes in epigenetic allele frequencies can be directly
113 attributed to the response to selection. To this aim, we developed a new approach that
114 consists in growing all the progeny as if every parental plant had been selected to be a
115 breeder and removing the progeny of counter selected parents *in silico*, which allowed us
116 to build strong confidence intervals and conduct independent selection trials on different
117 traits.

118 **Materials and Methods**

119 **Plant material**

120
121
122 Epigenetic Recombinant Inbred Lines (epiRILs) of *Arabidopsis thaliana* plants represent
123 one of the few study systems available to date to investigate epigenetic variation
124 associated with phenotypic trait variation, clearly identified as epiQTLs located on
125 chromosomes, in an almost fixed genomic background (13, 18, 20, 25). We obtained 123
126 epiRILs from the Institute Jean-Pierre Bourgin (IJPB - UMR1318 INRAE-AgroParisTech,
127 Versailles, France). These epiRILs were originally created by crossing two parents from
128 the same accession (*Columbia*, *Col*) that differed only at the *Decreased DNA Methylation*
129 *1* gene (*DDM1*), with one parent homozygous for the wild type *DDM1* allele (*Col-wt*),
130 and the other for the *ddm1-2* mutant allele (*Col-ddm1*) (Johannes et al. 2009). The latter is
131 characterized by a widely different DNA methylation profile background. These epiRILs
132 were obtained after six generations of single seed descent and harbor differentially
133 methylated regions (DMR) across the entire genome. These DMRs are stably inherited
134 over multiple generations and associated with substantial heritable variation in flowering
135 time and plant height (Cortijo et al. 2014, Johannes et al. 2009, Roux et al. 2011).
136 Genome-wide DNA methylation data are available for these 123 epiRILs as well as the
137 derived recombination map for 126 meiotically stable DMRs covering 81.9% of the
138 genome (Colomé-Tatché et al. 2012).
139

140 We built two independent samples of 60 different epiRILs by randomly choosing lines
141 without replacement out of the 123 lines available in order to create two independent
142 experimental populations of *Col-wt* epiRILs and replicate our selection experiment in
143 these two populations at different times. Population 1 included the following epiRILs
144 (Colomé-Tatché et al. 2012): 14, 20, 36, 52, 54, 55, 62, 64, 70, 92, 99, 108, 112, 144, 164,

145 166, 169, 183, 193, 195, 215, 216, 238, 257, 260, 275, 277, 297, 305, 326, 333, 340, 344,
146 350, 356, 361, 362, 368, 371, 393, 394, 400, 408, 432, 434, 437, 458, 466, 467, 473, 477,
147 480, 488, 492, 493, 497, 503, 539, 561 and 567. Population 2 included the following
148 epiRILs: 8, 11, 18, 24, 46, 53, 60, 69, 71, 73, 94, 95, 101, 114, 118, 137, 147, 148, 150,
149 159, 170, 172, 202, 208, 218, 222, 225, 229, 232, 244, 252, 258, 262, 276, 315, 323, 363,
150 366, 375, 391, 410, 425, 438, 439, 454, 471, 494, 495, 500, 506, 508, 523, 538, 556, 558,
151 559, 570, 572, 573 and 579.

152 Plant cultivation

153
154
155 The selection experiment required growing two successive generations of plants (G0 and
156 G1) for each of the two independent experimental populations of Col-wt epiRILs. G0 of
157 population 1 was grown in June 2018 and G1 in January 2019. G0 of population 2 was
158 grown January 2019 and G1 in August 2019. At G0, three plant replicates per epiRIL were
159 grown (n = 180 plants for each population). Ninety-eight percent of G0 plants survived
160 until flowering (n = 177 plants in population 1 and n = 177 plants in population 2). At G1,
161 three descendants of every G0 plant were grown (n = 531 plants per population). Because
162 of a final success rate of 96 %, 524 plants were available for population 1, and 521 plants
163 for population 2. All the plants were grown in a growth chamber under controlled long-
164 day conditions (day: 16 h – 20/22°C, night: 8 h – 16°C/18°C) with artificial light (120
165 mmol/m²s⁻¹), in 7x7x6.4 cm pots (Soparco, France) on a substrate composed of 80% BP
166 Substrate K716 (Klassman, France) and 20 % of vermiculite (Souflet Vigne, France).
167 About 5 to 10 seeds, equivalent in terms of lineage and parental plant, were sown in each
168 pot and seedlings were thinned out to retain a single plant per pot. Plant location in the
169 growth chamber and planting occurring over a few hours was fully randomized.
170 Comparisons were carried out under identical environmental conditions and the design
171 was fully randomized. G0 and G1 were grown 6 months apart in different seasons, but in
172 an underground growth chamber under fully controlled environmental conditions.

173 Plant phenotypic measurements

174
175
176 Four growth- and flowering-related phenotypic traits were recorded: Flowering time (in
177 number of days since sowing) and rosette diameter were recorded when the first flower
178 opened. Height at first silique and above-ground biomass were measured after plants
179 ceased flowering. Above-ground biomass was measured on an analytical high precision
180 lab balance (Sartorius, Germany) after plants were dried for 24h at 40°C in a laboratory
181 oven (France Etuves, France). These traits, and in particular growth traits, are not expected
182 to be fully independent. They were chosen because they present transversal and
183 comparative interest in ecological and evolutionary plant studies as they are usually
184 recorded in *A. thaliana* and other plant studies (Halbritter et al. 2018, Roux et al. 2011,
185 Herrera and Bazaga 2010).

186 Selection experiment

187
188
189 G0 populations formed the base generation upon which selection was conducted. In each
190 population, we ranked plants based on their trait values. We randomly selected plants from
191 this list without replacement to establish a control line. We then selected the top-ranked
192 and bottom-ranked of the remaining plants to establish high and low selection lines,
193 respectively. The high and low selection lines (HSL and LSL) respectively aim at
194 increasing and decreasing trait value in response to selection at the next generation. In

195 order to test the limits of the selection protocol, the selection experiment was conducted at
196 two selection intensities: weak selection, whereby each selection line (HSL and LSL)
197 consisted of 33% of the plants (60 plants out of 180) and strong selection, whereby ca.
198 11% of the plants were randomly sampled to build a control group and then the upper and
199 lower ca. 11% (20 plants out of 180) of the remaining plants were selected to build each
200 selection line (HSL and LSL). The sample size of control groups matched the size of
201 selection lines. G1 control group phenotypes, their similarity to G0, and their difference
202 with G1 selection lines, allows to control for confounding effects between the response to
203 selection and the effect of temporal changes in the controlled environment of the growth
204 chambers, which in itself is highly unlikely.

205
206 Selection experiments conducted on outcrossing plants require cross-pollinating different
207 sets of plants for the different traits under selection and the different selection lines, and
208 sometimes require isolating the plants forming control and selection lines before pollen
209 might be dispersed between groups. This was not necessary in our selection experiment
210 because Col-wt epiRILs reproduce almost exclusively by self-fertilization and were grown
211 in tubes that isolate flowering stems from each other. Here, we grew the progeny of every
212 G0 plant as if it had been chosen and multiplied (three descendants per parent) to be part
213 of the control, the HSL and LSL. This original protocol is challenging on the logistical
214 front because the number of plants cultivated in the growth chamber, and the number of
215 phenotypic records, grew exponentially between G0 and G1, but it has certain advantages.
216 It allowed us to establish HSL, LSL and control lines and estimate the response to
217 selection *in silico* based on the four different phenotypic traits after the two generations of
218 plants had been grown and measured *in vivo*.

219 220 Statistical analysis

221
222 The classical *in vivo* estimation of the response to selection can be biased because a
223 unique event of random sampling without replacement of a fixed set of individuals from
224 the base population is used to establish the control line. This protocol can potentially
225 produce a sample of values that is upward or downward biased by chance alone. The
226 remaining individuals available to establish the high and low selection lines are then
227 biased towards average values. Our original protocol allowed us to overcome this potential
228 bias although we applied the same logical random sampling of individuals without
229 replacement to build the control line. This is because we replicated 1000 times this
230 random sampling *in silico*. We built the corresponding distribution and estimated unbiased
231 estimates of the trait mean, Standard Error, and 95% Confidence Interval, for each line
232 (control group, HSL, LSL) and each selection intensity. We used the lack of overlap
233 between the 95% confidence intervals of phenotypic estimates to test for a significant
234 difference between high and low selection lines. This is a conservative approach that
235 provides information on the statistical effects (range, direction, strength and reliability)
236 that is not provided by p values 40. Scripts in R are publicly available on the ZENODO
237 repository: <https://doi.org/10.5281/zenodo.7603324>.

238 239 Phenotypic selection analysis

240
241 On the basis of the unbiased control and selection line values obtained on the basis of
242 1000 replicates, we estimated the response to selection (R). In our selection experiment,
243 the response to selection evaluates the change in trait mean obtained after one generation
244 of selection. R was typically estimated by the difference of mean trait value between the

245 selection line and control line in G1, which allows controlling for potential between-year
246 environmental variation of the mean phenotype (Walsh and Lynch 2018). We also
247 estimated the different parameters of the breeder's equation, $R = h^2S$, where S is the
248 selection differential and h^2 is the realized narrow-sense heritability (Lush 1937). The
249 selection differential evaluates the selection pressure imposed on the base population by
250 measuring the trait mean difference between the population and the selection line. The
251 selection differential (S) was typically estimated by the difference of mean trait value
252 between the selection and control line in G0 (Walsh and Lynch 2018). We estimated h^2 for
253 each trait on the basis of the available measurements of R and S. Caution must be taken
254 when interpreting the heritability parameter calculated by the breeder's equation in self-
255 fertilizing plants because it might in fact estimate broad-sense heritability (H^2).

256 257 Analysis of DNA methylation changes after selection

258
259 We tested *in silico* for changes in DNA methylation frequencies in response to phenotypic
260 selection conducted *in vivo*. We constructed heat maps to visualize the changes in mean
261 epi-allelic frequencies after selection for all the DMRs in both experimental populations
262 by using the package ggplot2 (Wickham 2016). To this aim, we estimated in G1 the mean
263 epi-allelic frequencies at each of the 126 DMRs, for each selection and control line, and at
264 each selection intensity and trait, on the basis of the 1000 resampled data sets. Epi-alleles
265 were defined as "0" when originating from the *DDMI* wild type Col-wt and "1" when
266 originating from the *ddm1-2* mutant col-ddm1. As a result, the more *ddm1-2* mutant epi-
267 alleles are present in a given line, the higher the frequency of DNA methylation.

268
269 We condensed the epi-allelic variation of each experimental population into a reduced
270 number of linearly independent dimensions by using a Principal Component Analysis.
271 PCA was done using the function dudi.pca of the ade4 package (Dray and Dufour 2007).
272 It summarized more than 90 % of epi-allelic variation in 16 dimensions for population 1
273 and in 17 dimensions for population 2. As a result, every plant had a coordinate on each of
274 those PCA dimensions. We estimated the average coordinate and its 95% CI based on the
275 1000 resampled datasets, for each PCA dimension, selection or control line, selection
276 intensity, and trait. We tested for significant differences in DNA methylation associated
277 with selection by comparing the 95% CIs of high and low lines in G1. We used the lack of
278 overlap between the 95% confidence intervals to test for a significant difference between
279 high and low selection lines. As mentioned above, this is a conservative approach
280 providing information on the statistical effects (range, direction, strength and reliability)
281 that is not provided by p values (Ho et al. 2019).

282 283 Analysis of molecular epigenomic integrity

284
285 Molecular epigenomic sequence data covering the whole genome were used to confirm
286 the integrity of the DMRs in the epiRILs that were used in the plant selection experiment.
287 Tissue sampling of 24 epiRILs (12 per population) and Col-wt was done 24 days after
288 germination. Two rosette leaves were sampled and put into a 2-ml tube that was directly
289 dropped into liquid nitrogen for instant freezing to ensure DNA conservation. Samples
290 were kept at -80°C . DNA extraction was performed using DNeasy kit (Qiagen, Germany)
291 following the manufacturer's protocol. Three replicates per EpiRIL were extracted
292 separately and DNA content was quantified by using a Qubit fluorometer with the dsDNA
293 HS Assay Kit (Invitrogen, USA). Quality was controlled using a NanoDrop
294 spectrophotometer (Thermo Fisher Scientific, USA). The three DNA replicates were then

295 pooled together and a control of the possible degradation of the DNA was performed by
296 using TapeStation automated electrophoresis (Agilent, USA). Enzymatic Methyl-
297 sequencing (EM-seq) was then conducted on these DNA samples by the IGen Seq
298 platform (ICM, Paris, France). EM-seq library was prepared with the NEBNext®
299 Enzymatic Methyl-seq (EM-seq™) (New England Biolabs, USA) following the
300 manufacturer's protocol. This method is known to provide accuracy and reliability at least
301 equivalent to that achieved by Whole Genome Bisulfite Sequencing (WGBS) with similar
302 global methylation levels indicating high resemblance in the overall detection of
303 methylated Cs (Feng et al. 2020, Hoppers et al. 2020). Only the bisulfite conversion step
304 differs between these two approaches; it is an enzymatic reaction in the EM-seq approach
305 while it is a sodium bisulfite chemical treatment in the WGBS approach. The first EM-seq
306 conversion step uses TET2 and an Oxidation Enhancer to protect modified cytosines from
307 downstream deamination. TET2 enzymatically oxidizes 5mC through a cascade reaction
308 into 5-carboxycytosine, protecting 5mC from deamination. The second enzymatic step
309 uses APOBEC, which deaminates cytosines but does not affect 5caC. As a result,
310 converted sequences are similar in the data generated by these two approaches and can be
311 analyzed in the same way. Sequencing was done on an Illumina Novaseq 6000 sequencer
312 (Illumina Inc, USA) with a flow cell S1 (300 cycles of 1600 million of reads (2*53 million
313 reads of 150 base pairs per sample).

314
315 We verified the quality of the sequence reads by using the FastQC software Version
316 0.11.9 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). We then trimmed 10
317 bp from each end of the paired-end reads and removed adapter sequences using Trim
318 Galore software Version 0.6.5
319 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were
320 subsequently aligned to the reference TAIR10 genome version of *Arabidopsis thaliana*
321 (Assembly: GCA_000001735.1) using BSMAP aligner Version 2.9 (Xi and Li 2009)
322 specifying an allowed minimal insert size of 40 pb and a maximal insert size of 2500 bp.
323 Only reads aligning uniquely to the reference genome were used for subsequent analyses.
324 Cytosines with an average coverage below 8 (and above 100) across each line were
325 removed to avoid a potential bias. Average genome coverage was 73.6 after filtering,
326 which is well above most previous studies on DNA methylation in plants (Schmid et al.
327 2018). Mapped reads were then used as input data in the BSMAP methylation caller
328 Version 1.0.0 implemented in the Galaxy interface (<http://usegalaxy.org>) (Goecks et al.
329 2010) to extract the methylation of each cytosine call and distinguish between the three
330 contexts of methylation: CG, CHG and CHH. We used the methylKit package Version
331 1.20.0 for DMR analysis (Akalın et al. 2012). False-positive methylation levels were
332 measured by calculating methylation levels in the unmethylated chloroplast genome. The
333 126 DMRs were identified in the epiRILs used in our experiment by calculating the mean
334 methylation ratio of each marker found by filtering genomic positions (chromosome and
335 start and stop bp) from the BSMAP methylation calling files with a custom script in R
336 made publicly available using the packages tidyverse and dplyr (Mailund 2019, Wickham
337 et al. 2019). These DMRs cover genomic regions of a size ranging from 200 base pairs to
338 11 kb. The distribution of the Bisulfite Sequence signals was verified for the 126 markers
339 by plotting the distribution of BS signals in boxplots of the WT-inherited (Methylated)
340 and the *ddm1*-inherited (Unmethylated) markers with a custom script in R made publicly
341 available on the ZENODO repository: <https://doi.org/10.5281/zenodo.7603324>.

342
343 For each DMR marker, percentages of methylated cytosines were compared between Col-
344 wt and each epiRIL by using Fisher's exact tests to confirm whether they were WT-

345 inherited (not significantly different from Col-wt) or ddm1-inherited (significantly
346 hypomethylated in comparison to Col-wt). To this aim, we used the methylKit package.
347 Parameters used to call for differentially methylated markers were defined as follows:
348 differential methylation > 25% and q value < 0.01 as established by other work on the
349 same species (Kim et al. 2017). This approach allowed us to identify statistical differences
350 between Col-wt and each epiRIL for the 126 differentially methylated regions (DMRs)
351 published in Colomé-Tatché et al. (2012), and then compare the DMRs found in our
352 experiment with those already published in the literature. To this aim, we calculated the
353 percentage of overlapping probe classification following the methods developed by
354 Colomé-Taché et al. (2012) between published epiRIL DMR data and our EM-seq data for
355 the plants used in our experiment. We considered percentages of similarity greater than
356 80% as a signal of very strong molecular correspondence, > 70% as strong, > 60% as
357 moderate, and < 50% as weak. We also evaluated the molecular correspondence between
358 the available WGBS data published for 8 epiRILs (60, 92, 150, 193, 202, 232, 260 and
359 480) (Bewick et al. 2016, Kooke et al. 2019, Lauss et al. 2018) and the EM-seq data of
360 these same epiRILs used in our experiment by comparing directly the bisulfited
361 sequencing signal. To this aim, we estimated the correlation between the 126 regions
362 methylation ratios obtained from both techniques by using a Kendall rank correlation
363 coefficient for each methylation context and for global methylation (cumulated
364 information for all methylation contexts). We also conducted a hierarchical cluster
365 analysis of the global methylation at the 126 DMRs of epiRILs and Col-wt, and then
366 compared the hierarchical clusters by using cophenetic correlation coefficients. We also
367 annotated the 126 DMRs using the PlantGDB database (Dong et al. 2004) and the
368 Assembly version: AtGDB TAIR9/10 v171 (<https://plantgdb.org/AtGDB/>).

369 Data and code availability

370
371
372 The EM-seq data used in this study are available on the NCBI bioproject online
373 repository. The names of the repository/repositories and accession number(s) can be found
374 below: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA773039>. Data and R protocols
375 used for the statistical analyses are available from the ZENODO Digital Repository
376 <https://doi.org/10.5281/zenodo.7603324>.

378 Results

379 DNA methylation and phenotypic traits vary

380
381
382 We assessed and confirmed that the two *A. thaliana* experimental populations harbored
383 DNA methylation polymorphism and trait variation in the *Columbia* wild type (Col-0)
384 genetic background. DNA methylation is the most studied epigenetic mark that influences
385 chromatin structure and the silencing of transposable elements (TE) and genes (19). Each
386 experimental population (n = 180 plants per population; Fig. 1A) consisted of a different
387 set of 60 distinct epigenetic recombinant inbred lines (epiRILs) out of the 123 available
388 epiRILs produced in *A. thaliana* (17, see materials and methods for details).

389 Molecular epigenomic sequence data covering the whole genome were used to confirm
390 the integrity of the DMRs in the epiRILs on the basis of three biological replicates per
391 epiRIL for 24 epiRILs. This molecular analysis confirmed that epiRILs differed as

392 expected by harboring DNA methylation polymorphism in the form of 126 Differentially
393 Methylated Regions (DMRs) previously identified and whose stable inheritance is
394 acknowledged (18, 20) (Figs. S1 and S2, Tables S1 and S2). Some of these 126 DMRs
395 function as epigenetic quantitative trait loci (13) that underlie the variation of the four
396 traits that we measured. Although we do not expect biomass, rosette diameter, flowering
397 time and height to be independent at the phenotypic level, the extent to which their
398 variation is associated with DNA methylation polymorphism was found to vary (20). In
399 the two artificial populations, we found 14% to 25% variation (estimated by the
400 coefficient of variation $100 \times \text{SE}/\text{mean}$) in above-ground dry biomass (average \pm SE in
401 populations 1 and 2, respectively; 0.64 ± 0.09 g and 0.60 ± 0.15 g), 12% to 23% variation
402 in rosette diameter (9.7 ± 1.12 cm and 7.7 ± 1.80 cm), 7% to 9% variation in flowering
403 time (27.3 ± 1.92 d and 28.3 ± 2.4 d), and 21% to 23% variation in height at first fruit
404 (12.1 ± 2.50 cm and 11.6 ± 2.69 cm). Thus, in accordance with the making of these
405 experimental populations as different subsets of epiRILs, we found phenotypic trait
406 variation and epigenetic variation in each of these two populations.

407 DNA methylation stability across generations

408 The potential lack of stability of DNA methylation across generations has raised the very
409 question of their evolutionary significance as a source of variation subject to selection (21-
410 24). Here, we conducted molecular correspondence analyses that confirmed the broad
411 stability of the 126 DMRs. Stability of DNA methylation polymorphism was assessed by
412 comparing the 126 DMRs (Colomé-Tatché et al. 2012) between reference epiRILs
413 established by Johannes et al. (2009) and the same epiRILs used in our selection
414 experiments, therefore after regeneration. This comparison confirmed their molecular
415 correspondence (Fig. S1). We used an Enzymatic Methyl-seq (EM-seq, 26) molecular
416 sequencing approach that provided us with the methylation status of the whole genome
417 cytosines of 24 epiRILs (12 per experimental population) with a high sequencing coverage
418 (see methods and Table S3). In plants, DNA methylation can occur in Cytosine-
419 phosphate-Guanine dinucleotides (CpG), CHG and CHH contexts (H = A, T or C) (19).
420 We found high molecular correspondence by using the overlapping probe classification
421 approach developed by Colomé-Tatché et al. (2012). This correspondence reached on
422 average 87%, ranging from 69% to 97%, for the CpG methylation context and on average
423 86%, ranging from 64% to 98%, for the CHG methylation context (Tables S1 and S2). We
424 did not investigate the CHH methylation context because it presents a low methylation
425 ratio, varies among cells, is often stochastic (27) and is less stable than CpG and CHG
426 contexts (28), and was not available in the literature (Colomé-Tatché et al. 2012).
427 Additionally, we compared the EM-seq data of 126 DMRs with the published WGBS data
428 available for eight epiRILs (29-31) (Fig. S2 and Table S4). We found an average
429 correlation of 0.65 (ranging between 0.49 and 0.79) for the CpG context, and an average
430 correlation of 0.73 (ranging between 0.64 and 0.79) for the CHG context. Furthermore, the
431 hierarchical clustering analysis indicated a high similarity of the methylation in epiRILs
432 sequenced by both EM-seq and WGBS approaches, with a high (0.89) cophenetic
433 correlation coefficient (Fig. S2). Our results were coherent with the knowledge that
434 hypomethylation is stable across generations in the CpG context (32) and with the stability
435 of hypomethylation in the CHG context where molecular mechanisms driving
436 modifications differ (33). Although the methylation status is known to be reversible in
437 DMRs (34, 35), we detected a very small number of changes in the 126 DMRs, illustrating
438 the broad stability across generations of DNA methylation polymorphism (Tables S1 and
439 S2).

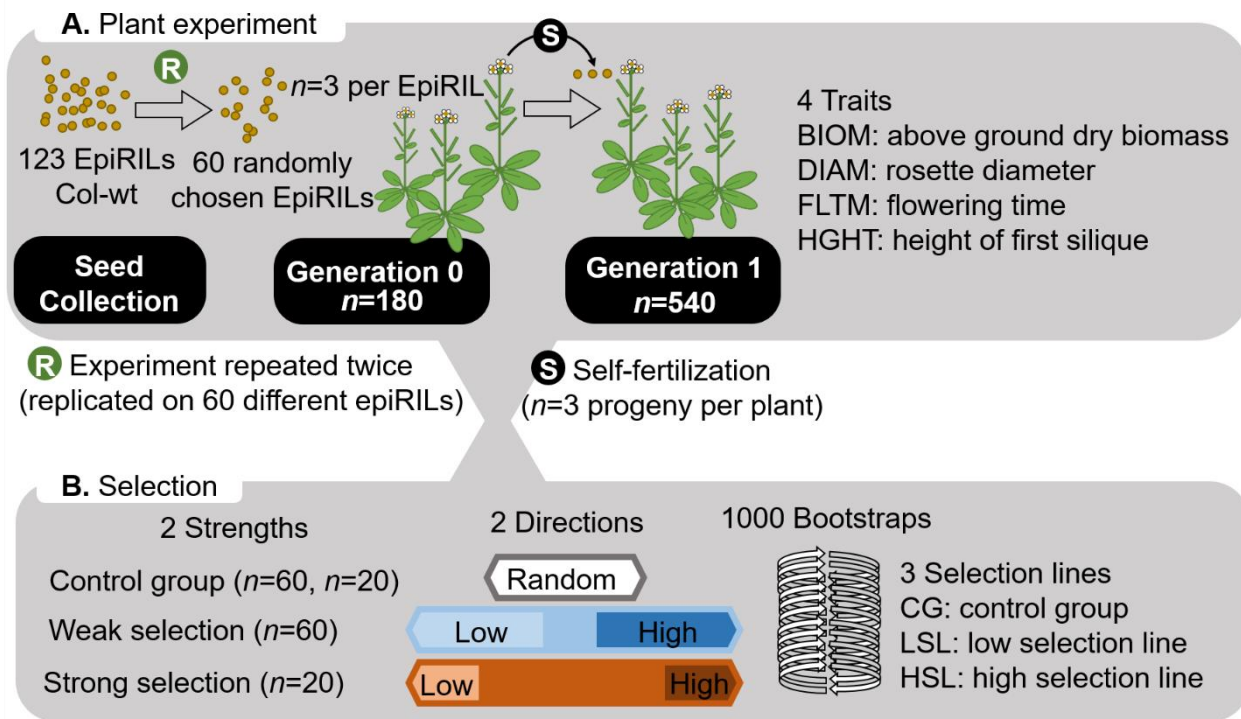
440 DNA hypomethylation can release TEs that can potentially impact gene expression and
441 compromise genome integrity (18, 35, 36). Novel TE insertions have been shown to
442 contribute to phenotypic variability in *A. thaliana* epiRIL populations (35, 37). However,
443 several studies on *A. thaliana* epiRILs however did not find that novel TE insertions
444 contributed to phenotypic variation (13, 18, 29, 35, 39). These studies qualified these
445 insertions as rare and random; most of them were found to be neutral (13, 23, 24, 39).
446 Most of these insertions are described as private, i.e., found in a single epiRIL, at the
447 hemizygous state (24). TEs are often considered not to affect estimates of phenotypic
448 variation and DNA methylation polymorphism estimated on the basis of multiple epiRILs
449 because their effect would be diluted in a population made of multiple epiRILs as a result
450 of DMRs differing by definition between epiRILs (13, 22, 39). Our *in silico* resampling
451 approach also warrants our results against any bias produced by a given epiRIL by making
452 any such potential effect infinitesimal. In the absence of resequencing approach, changes
453 in copy numbers of private TE insertions affecting a given epiRIL phenotype cannot be
454 strictly excluded. Our results showed that over the 126 DMRs, which included protein
455 coding genes, intergenic regions, and TEs, only 25 DMRs included potentially mobile TEs
456 (23) that could affect traits of interest (24) (Table S5). We therefore identified methylation
457 changes occurring at these 25 DMRs between the 24 epiRILs sequenced by the EM-seq
458 approach and the same epiRILs from the reference epiRIL data established by Johannes et
459 al. 2009 (18). Most of these DMRs did not change. Very few new hypomethylation
460 changes occurred in these DMRs, a mean of 0.2% (ranging from 0% to 4%) in CpG
461 context, and of 0.3% (ranging from 0% to 4%) in CHG context. On average 6.3% (ranging
462 from 0% to 32%) of these DMRs in the CG context and 9.6% (ranging from 0% to 36%)
463 of these DMRs in the CHG context reverted from a hypomethylated state to the wild-type
464 methylation state that is known to silence TEs and therefore keep the integrity of the
465 genome structure (Tables S1 and S2). One can nevertheless not exclude that high copy
466 numbers of private TEs might affect the phenotype of a given epiRIL even if as mentioned
467 above, the likelihood is low and the potential effect greatly diluted by our resampling
468 approach. Our selection experiment should therefore mostly target epiRIL phenotypic
469 variation directly caused by DMRs but caution is necessary because it may act
470 occasionally on genomic variants caused by high copy number TEs, which release would
471 be due to hypomethylation.

472 Phenotypic response to selection

473 Despite the absence of genetic variation, we tested whether each experimental population
474 responded to directional selection for higher and lower trait values. We artificially selected
475 plants based on their morphological (above ground biomass, rosette diameter, flowering
476 time) and reproductive trait (flowering time, height of first silique) values independently
477 in each population, and for each trait. We selected for upper and lower trait values in G0
478 plants and reproduced them to observe the response to selection in G1 plants. Typical
479 selection experiments require to use different sets of plants; one for each selection trial
480 based on a given trait. Here, we multiplied each G0 plant as if it had been positively
481 selected for (three descendants per parent). As a result, we obtained a large number of
482 plants in G1, so that we could build *in silico* the G1 dataset corresponding to a given
483 selection intensity based on a given trait. For each population (n = 180 plants per
484 population) and each trait, we randomly sampled a third of the plants to build a control
485 group. We then selected the remaining upper and the lower third (weak selection line; n =
486 60) of plants in terms of trait values to comprise high and low selection lines (HSL and
487 LSL, respectively), in each population and for each trait. In parallel, also between G0 and

488
489
490

G1, we applied a stronger selection pressure by randomly sampling ca. 11% of the plants to build a control group and selecting the upper and lower ca. 11% (strong selection line; $n = 20$) of the remaining plants (Fig. 1a, b).



491

Fig. 1. Selection experiment drives phenotypic trait changes. **A**, Plant experiment: Two sets of 60 different epiRILs (*Arabidopsis thaliana* Col-ddm1) were randomly assembled to build the two different experimental populations that were cultivated at different periods in the cultivation chamber. Seeds from the collection were grown to produce a buffer generation without selection and generate the seeds used to grow G0 plants. G1 plants – the progeny of G0 – were produced by self-fertilization. **B**, Selection design: At G0, a control group was made of randomly selected plants. Plants with the highest and lowest trait values were respectively used to build the high and low selection lines. Weak and strong selection lines were respectively made up of 33% and 11% of the population. Corresponding progenies were used to build control groups and selection lines in G1, with differences between G1 lines reflecting responses to selection.

503
504
505
506
507
508
509
510
511
512
513
514
515
516

In the progeny, we found significant divergence in response to selection between high and low selection lines in terms of rosette diameter and height in both populations (Table 1). We found trait changes in response to selection for size traits in both replicate populations of epiRILs, changes for flowering time, but not under every selection treatment, and no change for biomass. Statistical significance was established by lack of overlap between 95% Confidence Intervals (CIs) of trait means, which is a conservative approach (40). For each trait and selection strength, the 95% CIs were based on 1000 bootstraps generated in silico by 1000 independent random samples of the control group. For each randomly sampled control group, it was possible to establish the corresponding specific set of selection lines because we grew the progeny of every plant (Fig. 1A, see methods for details). As a result, potential biasing effects of random sampling, e.g., the weight of a given epiRIL, could not affect our analyses of selection lines and control groups. Divergence was also found in most cases for the flowering time, with the exception of population 2 where it was non-significant after weak selection. No significant response to

517 selection was detected for the above-ground biomass, in both populations at both selection
518 intensities. The largest trait change after selection was found for height at first fruit, with a
519 16% reduction observed in the LSL and a 13% increase in the HSL, but most trait changes
520 were around 3 to 8 percent (Table 1). Strong selection logically generated larger trait
521 changes in response to selection, but also more variable and therefore less predictable
522 selection responses (41), as illustrated by larger 95% CIs (Table S6). Heritability was
523 estimated as the proportion of phenotypic variation available for selection to act upon by
524 using the breeder's equation (42). It was absent for biomass, which had not changed in
525 response to selection, but relatively high for height at first fruit, rosette diameter and
526 flowering time (ranging from 0.13 to 0.63, Table 1), all of which changed after selection.
527 These results corroborate previous work on the inheritance of epiRIL trait variation (large
528 heritability of height and flowering time) or lack thereof (negligible heritability of dry
529 biomass) (20). We also found surprising results because rosette diameter, which was
530 originally expected to lack heritability based on previous work (20), showed some ability
531 to change under selection, which indicates heritability. Heritability is the heritable part of
532 phenotypic variation that is available to selection. Yet it is not rare to observe discrepancy
533 between heritability estimates based on the phenotypic variation of related individuals and
534 those based on the actual response to selection (43, 44).

535 **Table 1. Selection differential (S), response to selection (R) and heritability (h²) for**
536 **every selection line.**

Pop	Trait	Weak selection						Strong selection					
		LSL			HSL			LSL			HSL		
		S	R	H ²	S	R	H ²	S	R	H ²	S	R	H ²
1	Biomass	0.07	NS	N/A	0.07	NS	N/A	0.16	NS	N/A	0.14	NS	N/A
	Diameter	0.87	0.42 (5%)	0.48	0.89	0.41 (5%)	0.46	1.90	0.59 (7%)	0.31	1.74	0.66 (8%)	0.38
	Flowering time	1.36	0.85 (3%)	0.63	1.34	0.80 (3%)	0.60	2.65	1.28 (5%)	0.48	3.41	1.81 (6%)	0.53
	Height	1.85	0.41 (4%)	0.22	1.85	0.41 (4%)	0.22	3.80	1.28 (11%)	0.34	3.94	1.50 (13%)	0.38
2	Biomass	0.11	NS	N/A	0.11	NS	N/A	0.24	NS	N/A	0.23	NS	N/A
	Diameter	0.97	0.55 (6%)	0.57	1.69	0.72 (8%)	0.42	1.76	0.67 (8%)	0.38	3.90	1.13 (13%)	0.29
	Flowering time	1.63	NS	N/A	1.72	NS	N/A	2.77	0.87 (3%)	0.31	4.88	0.62 (2%)	0.13
		2.11	0.99	0.47	1.91	0.90	0.47	4.05	2.04	0.50	4.60	1.47	0.32

	Height	(8%)	(7%)	(16%)	(11%)
--	--------	------	------	-------	-------

537

538 Weak selection: selection by truncation of the upper and the lower third (n = 60) of
539 *Arabidopsis thaliana* plants ranked by trait values to compose high and low selection lines
540 (HSL and LSL, respectively). Strong selection: same method but selecting by truncation
541 the upper and lower ca. 11% (n = 20) of plants. R is presented as trait change in trait units
542 (g, cm, d) and percentage change. Narrow-sense heritability (h^2) calculated by the
543 breeder's equation is in fact broad-sense heritability (H^2) in self-fertilizing plants.
544 Selection was considered significant following a conservative approach avoiding sampling
545 bias; when the 95% Confidence Intervals of average trait values estimated on the basis of
546 1000 bootstraps did not overlap between upper and lower selection lines (Table S6).
547 Biomass: above-ground dry biomass (g). Diameter: rosette diameter (cm). Flowering time
548 (d). Height: Height at first fruit (cm)

549 Epigenetic signature of selection

550 We tested whether responses to selection are mirrored by changes in epigenetic allele
551 frequencies at DMRs. After selection, we observed modifications of these frequencies (see
552 Fig. 2 presenting the average frequencies based on 1000 *in silico* replicates for all
553 selection and control lines in each experimental population). Estimating these frequencies
554 on the basis of 1000 bootstraps excluded the potential bias caused by the weight of a given
555 epiRIL in any given selection line or control group. These frequencies reflected divergent
556 epigenetic changes in the progeny of high and low selection lines in each population,
557 notably under strong selection (Fig. 2). Since sets of epiRILs differed between
558 populations, selection lines and control groups, it is therefore legitimate to observe
559 different frequencies for a given trait or chromosome region between populations, and to
560 denote the signature of selection depicted by a similar enrichment of epigenetic alleles on
561 the heatmap (Fig 2).

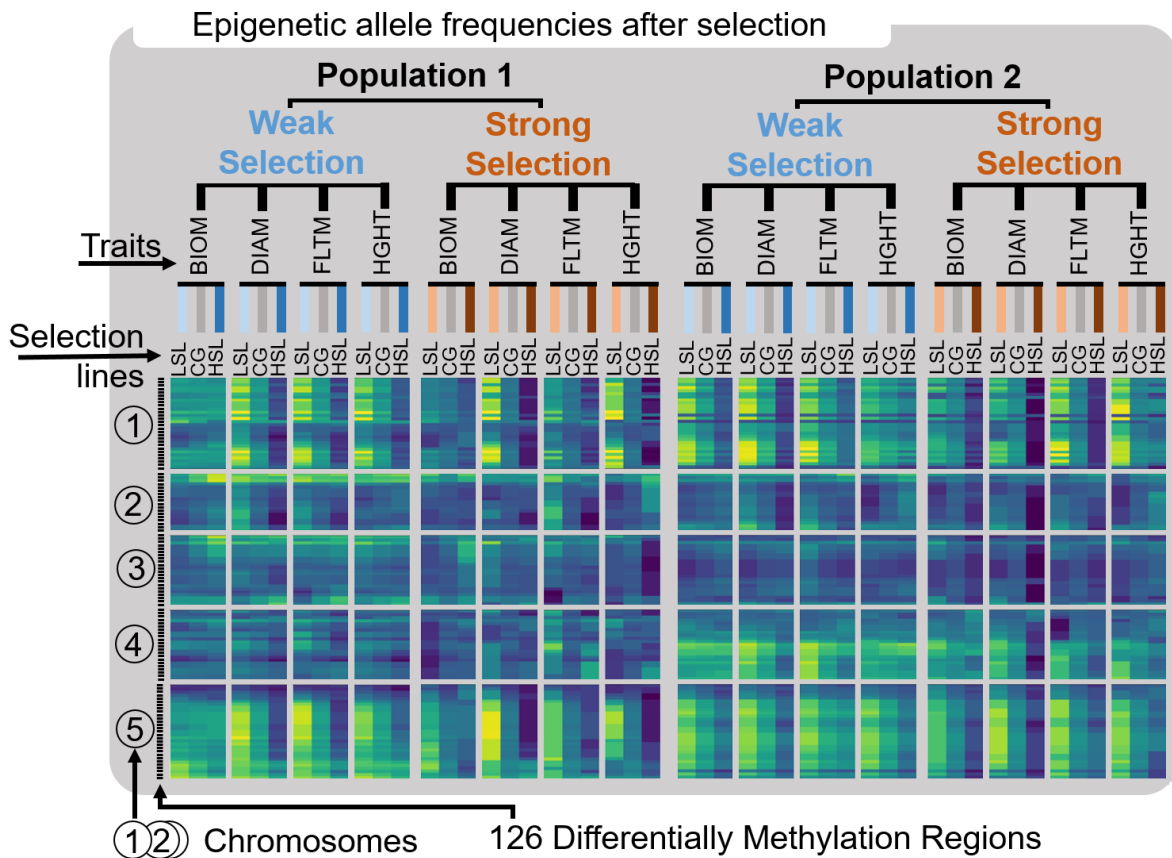
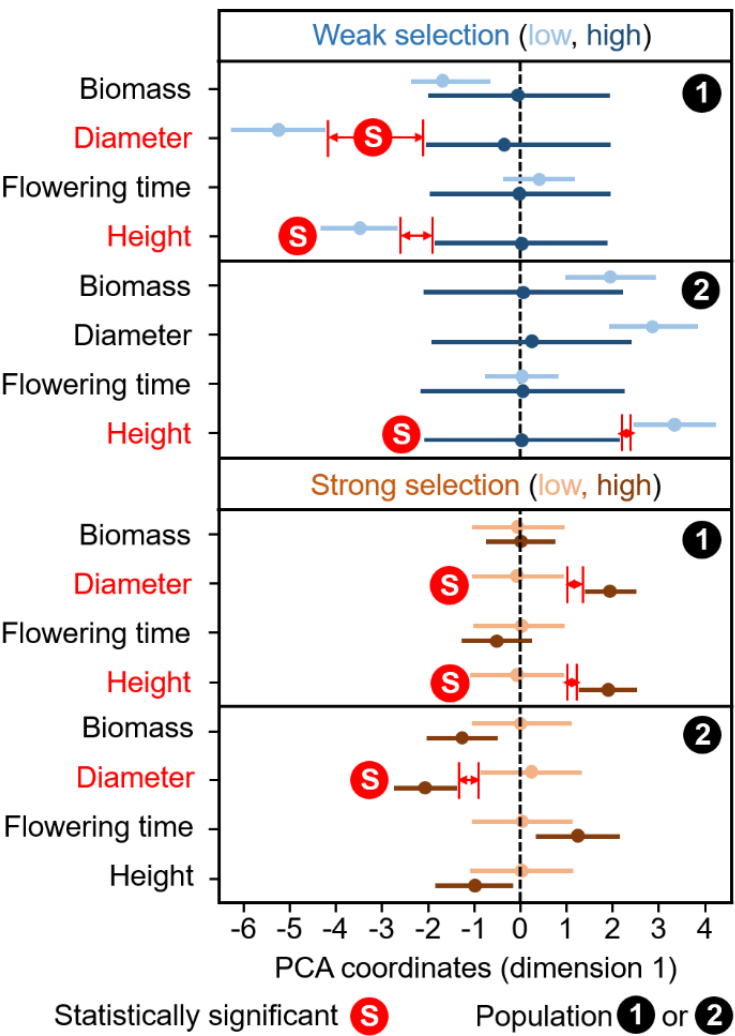


Fig. 2. Selection experiment drives changes in epigenetic frequencies. Changes in epigenetic allele frequencies after selection: The heat map presents DMR frequencies in G1 selection lines (LSL: low selection line and HSL: high selection line) and control groups (CG) for each of the four traits (BIOM: above ground dry biomass, DIAM: rosette diameter, FLTM: flowering time, HGHT: height of first silique) in populations 1 and 2. DMRs location on the five chromosomes are labelled on the left-hand vertical axis. Methylation frequencies of DMRs are represented by a colored band, with a gradient from blue (wild-type methylation status) to yellow (fixation of *ddm1-2* mutant methylation status). In each panel, DMR frequencies on a given chromosome are presented for the low, control and high selection lines. Each column corresponds to a trait. Frequencies were estimated on the basis of 1000 bootstraps.

We used a Principal Component Analysis (PCA) to reduce the dimensionality of epigenomic data and summarize its variation distributed over the whole genome into 16 independent variables that can be statistically analyzed. We found significant epigenetic divergence after selection on rosette diameter and height at first fruit in both populations (lack of overlap between 95% confidence intervals for PCA coordinates on the first dimension, representing 18% and 22% of the global DMR variation in populations 1 and 2, respectively, Fig. 3). Thus, the divergent selection for rosette wider and narrower rosette diameters led to opposite changes in the frequency of cytosine methylations in the same genetic background. A similar epigenetic response to selection was observed after divergent selection for smaller and higher plant height at first fruit. Phenotypic selection for either increased or decreased above-ground biomass did not trigger a significant phenotypic change in response to selection. It was therefore not surprising to observe the absence of significant epigenetic divergence in this trait between selection lines. Interestingly, selection for earlier and later flowering plants, which had resulted in

588
589
590
591
592
593
594

moderate responses to selection in population 1 and almost no response in population 2, did not drive statistically significant epigenetic changes (Fig. 3). This lack of epigenetic response may reflect the conservative nature of our approach or the possibility that the epigenetic architecture of flowering time consists of one to few loci whose variation was not captured by our statistical approach. Our study in plants therefore provides direct experimental evidence for heritable epigenetic changes in response to selection based on size.



595

Fig. 3. Epigenetic divergence or its lack thereof between selection treatments. Epigenetic divergence for each *Arabidopsis thaliana* epiRIL population (pop 1 and 2), each trait, each selection strength (n = 60 for each line under weak selection, n = 20 for each line under strong selection), and line (high and low selection lines) of plants. Principal Component Analysis coordinates for the first dimension (PC1), which synthesizes 18% and 22% of the global DMR variation in populations 1 and 2 respectively, are presented with their 95% confidence intervals. Divergence was considered significant following a conservative approach avoiding sampling bias: when 95% Confidence Intervals of PC1 values estimated on the basis of 1000 bootstraps did not overlap between upper and lower selection lines. Significant epigenetic divergent responses to selection are shown in red.

607 **Discussion**

608 Theory predicts that selection acts on transgenerational epigenetic variation presumably
609 without a contribution of genetic variation (5, 7). Evidence for epigenetic divergence
610 between populations sharing different histories of adaptation suggests such shaping of
611 epigenetic trait variation by selection in both plants and animals (17, 45-51). Our selection
612 experiment contributed to advancing this knowledge by providing a test confirming this
613 expectation in real time over a single generation of selection, through the comparison of
614 epiRIL experimental populations of *A. thaliana* surveyed before and after selection.
615

616 Our study did not aim to infer which causal epigenetic variants were responsible
617 for which phenotypic trait value, but rather to test for an epigenetic response to selection.
618 It is nevertheless interesting to note that the range of artificially induced epigenetic trait
619 variation in *A. thaliana* epiRILs is similar to the heritable fitness-related trait variation
620 observed in divergent natural populations (20). Although the DNA methylation
621 polymorphism of epiRILs is larger than what can be observed in natural populations
622 (Becker et al., 2011), our results provide us with evidence that a range of trait epigenetic
623 heritabilities can be used to predict a range phenotypic changes in response to selection.
624 Furthermore, as illustrated by Figure 2, LSLs tend to have higher DNA methylation state
625 variability than HSLs, which appears to be mostly due to Chromosomes 1 and 5. In *A.*
626 *thaliana* epiRILs, an increasing DMR frequency represents an increase in the proportion
627 of *ddm1-2* mutant epigenetic alleles in the population. LSLs therefore harbor more
628 hypomethylated loci than control and HSLs harbor more hypermethylated loci in
629 comparison to control. This finding suggests that our selection for bigger epiRIL plants
630 that flower later selected against the hypomethylated nature of epiRILs that was
631 artificially created by using the *ddm1* mutation. The artificially induced global
632 hypomethylation of the *A. thaliana* epiRIL genome produced smaller plants flowering
633 earlier. In fact, we selected for epigenomes and phenotypes that are closer to wild-type
634 plants.
635

636 Our experimental results show that selection can act on stable transgenerational epigenetic
637 variation presumably without a contribution of genetic variation. Our findings therefore
638 imply the contribution of DNA methylation polymorphism to short-term adaptation in
639 natural *A. thaliana* populations. However, one of the limitations of our experimental
640 setting is that it does not allow us to predict whether the epigenetic response to selection
641 plays a predominant role in nature, where selection is a much more complex mechanism.
642 In nature, selection will not target independently genetic and epigenetic components of
643 fitness-related traits (43). Natural selection will affect simultaneously all components of
644 fitness variation, which encompasses tightly linked genetic and epigenetic variation, some
645 reversible methylation patterns and hypomethylation driven release of TEs modifying the
646 integrity of the genome structure.
647

648 The scope for replication in other species of our proof of concept experimental approach is
649 limited. This is because EpiRIL biological material is only available in very few species
650 (20, 25). Heritable epigenetic variation for traits influencing fitness under environmental
651 changes nevertheless exists in both plants and animals (18, 45-52), suggesting that our
652 finding may be a common feature in many species. Other species harboring
653 transgenerational epigenetic variation for fitness-related traits may benefit from this
654 additional source of short-term adaptive or breeding potential (53, 54).
655
656
657

659 **References**

- 661 A. Akalin, M. Kormaksson, S. Li, F. E. Garrett-Bakelman, M. E. Figueroa, A. Melnick, C. E.
 662 Mason, methylKit: a comprehensive R package for the analysis of genome-wide DNA
 663 methylation profiles. *Genome Biol.* **13**, R87 (2012).
- 664 M. R. Baerwald, M. H. Meek, M. R. Stephens, R. P. Nagarajan, A. M. Goodbla, K. M. H. Tomalty,
 665 G. H. Thorgaard, B. May, K. M. Nichols. Migration-related phenotypic divergence is
 666 associated with epigenetic modifications in rainbow trout. *Mol. Ecol.* **25**, 1785-1800 (2016).
- 667 N. H. Barton, M. Turelli, Evolutionary quantitative genetics: how little do we know? *Annu. Rev.*
 668 *Genet.* **23**, 337-370 (1989).
- 669 C. Becker, J. Hagemann, J. Müller, D. Koenig, O. Stegle, K. Borgwardt, D. Weigel, Spontaneous
 670 epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480**, 245–9 (2011).
- 671 A. J. Bewick, L. Ji, C. E. Niederhuth, E.-M. Willing, B. T. Hofmeister, X. Shi, L. Wang, Z. Lu, N.
 672 A. Rohr, B. Hartwig, C. Kiefer, R. B. Deal, J. Schmutz, J. Grimwood, H. Stroud, S. E.
 673 Jacobsen, K. Schneeberger, X. Zhang, R. J. Schmitz, On the origin and evolutionary
 674 consequences of gene body DNA methylation. *Proc. Natl. Acad. Sci. USA* **113**, 9111-9116
 675 (2016).
- 676 R. Bonduriansky, T. Day, Nongenetic inheritance and its evolutionary implications. *Annu. Rev.*
 677 *Ecol. Evol. Syst.* **40**, 103–125 (2009).
- 678 D. Charlesworth, N. H. Barton, B. Charlesworth, The sources of adaptive variation. *Proc. R. Soc.*
 679 *B.* **284**, 20162864 (2017).
- 680 M. Colomé-Tatché, S. Cortijo, R. Wardenaar, L. Morgado, B. Lahouze, A. Sarazin, M. Etcheverry,
 681 A. Martin, S. Feng, E. Duvernois-Berthet, K. Labadie, P. Wincker, S. E. Jacobsen, R. C.
 682 Jansen, V. Colot, F. Johannes, Features of the *Arabidopsis* recombination landscape resulting
 683 from the combined loss of sequence variation and DNA methylation. *Proc. Natl. Acad. Sci.*
 684 *USA* **109**, 16240-16245 (2012).
- 685 S. Cortijo, R. Wardenaar, M. Colomé-Tatché, A. Gilly, M. Etcheverry, K. Labadie, E. Caillieux, F.
 686 Hospital, J.-M. Aury, P. Wincker, F. Roudier, R. C. Jansen, V. Colot, F. Johannes. Mapping
 687 the epigenetic basis of complex traits. *Science* **343**, 1145–8 (2014).
- 688 É. Danchin, A. Charmantier, F. A. Champagne, A. Mesoudi, B. Pujol, S. Blanchet, Beyond DNA:
 689 integrating inclusive inheritance into an extended theory of evolution. *Nat. Rev. Genet.* **12**,
 690 475–86 (2011).
- 691 Q. Dong, S. D. Schlueter, V. Brendel, PlantGDB, plant genome database and analysis
 692 tools. *Nucleic Acids Res.* **32**, D354-D359 (2004).
- 693 S. Dray, A.-B. Dufour, The ade4 package: implementing the duality diagram for ecologists. *J. Stat.*
 694 *Soft.* **22**, 1-20 (2007).
- 695 S. Feng, Z. Zhong, M. Wang, S. E. Jacobsen, Efficient and accurate determination of genome-wide
 696 DNA methylation patterns in *Arabidopsis thaliana* with enzymatic methyl
 697 sequencing. *Epigenetics Chromatin* **13**, 42 (2020).
- 698 J. Goecks, A. Nekrutenko, J. Taylor, The Galaxy Team, Galaxy: a comprehensive approach for
 699 supporting accessible, reproducible, and transparent computational research in the life
 700 sciences. *Genome Biol.* **11**, R86 (2010).

- 701 O. González-Recio, M. A. Toro, A. Bach, Past, present, and future of epigenetics applied to
702 livestock breeding. *Front. Genet.* **6**, 305 (2015).
- 703 Q. Gouil, D. C. Baulcombe, DNA Methylation signatures of the plant
704 chromomethyltransferases. *PLoS Genet.* **12**, e1006526 (2016).
- 705 M. P. Groot, N. Wagemaker, N. J. Ouborg, K. J. F. Verhoeven, P. Vergeer, Epigenetic population
706 differentiation in field- and common garden-grown. *Ecol. Evol.* **8**, 3505-3517 (2018).
- 707 A. H. Halbritter, S. Fior, I. Keller, R. Billeter, P. J. Edwards, R. Holderegger, S. Karrenberg, A. R.
708 Pluess, A. Widmer, J. M. Alexander, Trait differentiation and adaptation of plants along
709 elevation gradients. *J. Evol. Biol.* **31**, 784–800 (2018).
- 710 K. D. Harris, A. Zemach, Contiguous and stochastic CHH methylation patterns of plant DRM2 and
711 CMT2 revealed by single-read methylome analysis. *Genome Biol.* **21**, 194 (2020). E. Jablonka,
712 M. J. Lamb, A. Zeligowski, *Evolution In Four Dimensions. Genetic, Epigenetic, Behavioral,
713 And Symbolic Variation In The History Of Life* (MIT press, Cambridge, MA, 2014).
- 714 C. M. Herrera, P. Bazaga, Epigenetic differentiation and relationship to adaptive genetic
715 divergence in discrete populations of the violet *Viola cazorlensis*. *New. Phytol.* **187**, 867-876
716 (2010).
- 717 J. Ho, T. Tumkaya, S. Aryal, H. Choi, A. Claridge-Chang, Moving beyond P values: data analysis
718 with estimation graphics. *Nat. Methods* **16**, 565-566 (2019).
- 719 A. Hoppers, L. Williams, V. K. C. Ponnaluri, B. Sexton, L. Saleh, M. Campbell, K. Marks, M.
720 Samaranyake, L. Ettwiller, S. Guan, H. Church, B. Langhorst, Z. Sun, T. C. Evans Jr., R.
721 Vaisvila, E. Dimalanta, F. Stewart, Enzymatic methyl-seq: next generation methylomes. *J.*
722 *Biomol. Tech.* **31**, S15 (2020).
- 723 F. Johannes, E. Porcher, F. K. Teixeira, V. Saliba-Colombani, M. Simon, N. Agier, A. Bulski, J.
724 Albuissou, F. Heredia, P. Audigier, D. Bouchez, C. Dillmann, P. Guerche, F. Hospital, V.
725 Colot, Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS*
726 *Genet.* **5**, e1000530 (2009).
- 727 F. Johannes, M. Colomé-Tatché, Quantitative epigenetics through epigenomic perturbation of
728 isogenic lines. *Genetics* **188**, 215-227 (2011).
- 729 K. M. Johnson, M. W. Kelly, Population epigenetic divergence exceeds genetic divergence in the
730 Eastern oyster. *Evol. Appl.* **13**, 945-959 (2020).
- 731 I. Kakoulidou, E. V. Avramidou, M. Baránek, S. Brunel-Muguet, S. Farrona, F. Johannes, E.
732 Kaiserli, M. Lieberman-Lazarovich, F. Martinelli, V. Mladenov, P. S. Testillano, V.
733 Vassileva, S. Maury,. Epigenetics for crop improvement in times of global change. *Biology*
734 *(Basel)* **10**, 766 (2021).
- 735 G. Kim, C. R. Clarke, H. Larose, H. T. Tran, D. C. Haak, L. Zhang, S. Askew, J. Barney, J. H.
736 Westwood, Herbicide injury induces DNA methylome alterations in *Arabidopsis*. *Peer J.* **5**,
737 e3560 (2017).
- 738 R. Kooke, F. Johannes, R. Wardenaar, F. Becker, M. Etcheverry, V. Colot, D. Vreugdenhil, J. J. B.
739 Keurentjes, Epigenetic basis of morphological variation and phenotypic plasticity in
740 *Arabidopsis thaliana*. *Plant Cell* **27**, 337-348 (2015).
- 741 R. Kooke, L. Morgado, F. Becker, H. van Eekelen, R. Hazarika, Q. Zheng, R. C. H. de Vos, F.
742 Johannes, J. J. B. Keurentjes, Epigenetic mapping of the *Arabidopsis* metabolome reveals
743 mediators of the epigenotype-phenotype map. *Genome Res.* **29**, 96-106 (2019).

- 744 I. Kronholm, A. Bassett, D. Baulcombe, S. Collins, Epigenetic and Genetic Contributions to
745 Adaptation in *Chlamydomonas*. *Mol. Biol. Evol.* **34**, 2285–2306 (2017).
- 746 M. Latutrie, D. Gourcilleau, B. Pujol, Epigenetic variation for agronomic improvement: an
747 opportunity for vegetatively propagated crops. *Am. J. Bot.* **106**, 1281–1284 (2019).
- 748 K. Lauss, R. Wardenaar, R. Oka, M. H. A. van Hulten, V. Guryev, J. J. B. Keurentjes, M. Stam, F.
749 Johannes. Parental DNA Methylation states are associated with heterosis in epigenetic
750 hybrids. *Plant. Physiol.* **176**, 1627–1645 (2018).
- 751 K. D. Lemmen, K. J. Verhoeven, S. A. Declerck, Experimental evidence of rapid heritable
752 adaptation in the absence of initial standing genetic variation. *Funct. Ecol.* **36**, 226–238
753 (2022).
- 754 J. L. Lush, *Animal Breeding Plans* (Iowa State Collegiate Press, Ames, 1937).
- 755 T. Mailund, *Manipulating Data Frames: dplyr. R Data Science Quick Reference* (Apress,
756 Berkeley, CA, 2019).
- 757 O. Mathieu, J. Reinders, M. Caikovski, C. Smathajitt, J. Paszkowski, Transgenerational stability of
758 the *Arabidopsis* epigenome is coordinated by CG methylation. *Cell* **130**, 851–862 (2007).
- 759 M. Mirouze, J. Reinders, E. Bucher, T. Nishimura, K. Schneeberger, S. Ossowski, J. Cao, D.
760 Weigel, J. Paszkowski, O. Mathieu, Selective epigenetic control of retrotransposition in
761 *Arabidopsis*. *Nature* **461**, 427–430 (2009).
- 762 A. Muyle, J. Ross-Ibarra, D. K. Seymour, B. S. Gaut, Gene body methylation is under selection in
763 *Arabidopsis thaliana*. *Genetics* **218**, iyab061 (2021).
- 764 C. E. Niederhuth, R. J. Schmitz, Covering your bases: inheritance of DNA methylation in plant
765 genomes. *Mol. Plant.* **7**, 472–480 (2014).
- 766 J. M. Noshay, N. M. Springer, Stories that can't be told by SNPs; DNA methylation variation in
767 plant populations. *Curr. Opin. Plant Biol.* **61**, 101989 (2021).
- 768 M. Ong-Abdullah, J. M. Ordway, N. Jiang, S.-E. Ooi, S.-Y. Kok, N. Sarpan, N. Azimi, A. T.
769 Hashim, Z. Ishak, S. K. Rosli, F. A. Malike, N. A. Abu Bakar, M. Marjuni, N. Abdullah, Z.
770 Yaakub, M. Din Amiruddin, R. Nookiah, R. Singh, E.-T. L. Low, K.-L. Chan, N. Azizi, S. W.
771 Smith, B. Bacher, M. A. Budiman, A. Van Brunt, C. Wischmeyer, M. Beil, M. Hogan, N.
772 Lakey, C.-C. Lim, X. Arulandoo, C.-K. Wong, C.-N. Choo, W.-C. Wong, Y.-Y. Kwan, S. S.
773 R. S. Alwee, R. Sambanthamurthi, R. A. Martienssen, Loss of Karma transposon methylation
774 underlies the mantled somaclonal variant of oil palm. *Nature* **525**, 533–537 (2015).
- 775 S. Ossowski, K. Schneeberger, J. I. Lucas-Lledó, N. Warthmann, R. M. Clark, R. G. Shaw, D.
776 Weigel, M. Lynch, The rate and molecular spectrum of spontaneous mutations in *Arabidopsis*
777 *thaliana*. *Science* **327**, 92–94 (2010).
- 778 B. Pujol, S. Blanchet, A. Charmantier, E. Danchin, B. Facon, P. Marrot, F. Roux, I. Scotti, C.
779 Téplitsky, C. E. Thomson, I. S. Winney. The missing response to selection in the wild. *Trends*
780 *Ecol. Evol.* **33**, 337–346 (2018).
- 781 L. Quadrana, V. Colot, Plant transgenerational epigenetics. *Annu. Rev. Genet.* **50**, 467–491 (2016).
- 782 L. Quadrana, M. Etcheverry, A. Gilly, E. Caillieux, M.-A. Madoui, J. Guy, A. Bortolini Silveira, S.
783 Engelen, V. Baillet, P. Wincker, J.-M. Aury, V. Colot, Transposition favors the generation of
784 large effect mutations that may facilitate rapid adaption. *Nat. Commun.* **10**, 3421 (2019).
- 785 R Core Team, “R: A language and environment for statistical computing. R Foundation for
786 Statistical Computing” (2023; <https://www.R-project.org/>)

- 787 J. Reinders, B. B. H. Wulff, M. Mirouze, A. Mari-Ordóñez, M. Dapp, W. Rozhon, E. Bucher, G.
788 Theiler, J. Paszkowski, Compromised stability of DNA methylation and transposon
789 immobilization in mosaic *Arabidopsis* epigenomes. *Genes Dev.* **23**, 939-950 (2009).
- 790 C. L. Richards, C. Alonso, C. Becker, O. Bossdorf, E. Bucher, M. Colomé-Tatché, W. Durka, J.
791 Engelhardt, B. Gaspar, A. Gogol-Döring, I. Grosse, T. P. van Gurp, K. Heer, I. Kronholm, C.
792 Lampei, V. Latzel, M. Mirouze, L. Opgenoorth, O. Paun, S. J. Prohaska, S. A. Rensing, P. F.
793 Stadler, E. Trucchi, K. Ullrich, K. J. F. Verhoeven, Ecological plant epigenetics: Evidence
794 from model and non-model species, and the way forward. *Ecol. Lett.* **20**, 1576–1590 (2017).
- 795 F. Roux, M. Colomé-Tatché, C. Edelist, R. Wardenaar, P. Guerche, F. Hospital, V. Colot, R. C.
796 Jansen, F. Johannes, Genome-wide epigenetic perturbation jump-starts patterns of heritable
797 variation found in nature. *Genetics* **188**, 1015–1017 (2011).
- 798 M. W. Schmid, C. Heichinger, D. Coman Schmid, D. Guthörl, V. Gagliardini, R. Bruggmann, S.
799 Aluri, C. Aquino, B. Schmid, L. A. Turnbull, U. Grossniklaus, Contribution of epigenetic
800 variation to adaptation in *Arabidopsis*. *Nat. Commun.* **9**, 4446 (2018).
- 801 R. J. Schmitz, M. D. Schultz, M. G. Lewsey, R. C. O'Malley, M. A. Urich, O. Libiger, N. J.
802 Schork, J. R. Ecker, Transgenerational epigenetic instability is a source of novel methylation
803 variants. *Science* **334**, 369-373 (2011).
- 804 R. K. Slotkin, R. Martienssen, Transposable elements and the epigenetic regulation of the
805 genome. *Nat. Rev. Genet.* **8**, 272-285 (2007).
- 806 D. Stajic, L. Perfeito, L. E. T. Jansen, Epigenetic gene silencing alters the mechanisms and rate of
807 evolutionary adaptation. *Nat. Ecol. Evol.* **3**, 491–498 (2019).
- 808 D. Stajic, L. E. T. Jansen, Empirical evidence for epigenetic inheritance driving evolutionary
809 adaptation. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **376**, 20200121 (2021).
- 810 C. E. Thomson, I. S. Winney, O. C. Salles, B. Pujol, A guide to using a multiple-matrix animal
811 model to disentangle genetic and nongenetic causes of phenotypic variance. *PLoS*
812 *ONE* **13**, e0197720 (2018).
- 813 K. J. F. Verhoeven, B. M. vonHoldt, V. L. Sork, Epigenetics in ecology and evolution: what we
814 know and what we need to know. *Mol. Ecol.* **25**, 1631–1638 (2016).
- 815 G. Vernaz, M. Malinsky, H. Svardal, M. Du, A. M. Tyers, M. E. Santos, R. Durbin, M. J. Genner,
816 G. F. Turner, E. A. Miska, Mapping epigenetic divergence in the massive radiation of Lake
817 Malawi cichlid fishes. *Nat. Commun.* **12**, 5870 (2021).
- 818 G. Vogt, Epigenetic variation in animal populations: Sources, extent, phenotypic implications, and
819 ecological and evolutionary relevance. *J. Biosci.* **46**, 24 (2021).
- 820 B. Walsh, M. Lynch, *Evolution And Selection Of Quantitative Traits* (Oxford University Press,
821 Oxford, 2018).
- 822 H. Wickham, *ggplot2: Elegant Graphics For Data Analysis* (Springer-Verlag, New York, 2016).
- 823 H. Wickham, M. Averick, J. Bryan, W. Chang, L. D'Agostino McGowan, R. François, G.
824 Golemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache,
825 K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C.
826 Wilke, K. Woo, H. Yutani, Welcome to the Tidyverse. *J. Open Source Soft.* **4**, 1686 (2019).
- 827 A. J. Wilson, A. Rambaut, Breeding racehorses: what price good genes? *Biol. Lett.* **4**, 173–175
828 (2008).

- 829 Y. Xi, W. Li, BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinform.* **10**,
830 232 (2009).
- 831 H. Zhang, Z. Lang, J.-K. Zhu, Dynamics and function of DNA methylation in plants. *Nat. Rev.*
832 *Mol. Cell. Biol.* **19**, 489–506 (2018).
- 833 Y. Zhang, V. Latzel, M. Fischer, O. Bossdorf, Understanding the evolutionary potential of
834 epigenetic variation: a comparison of heritable phenotypic variation in epiRILs, RILs, and
835 natural ecotypes of *Arabidopsis thaliana*. *Heredity* **121**, 257–265 (2018).
- 836 Y. Y. Zhang, M. Fischer, V. Colot, O. Bossdorf, Epigenetic variation creates potential for evolution
837 of plant phenotypic plasticity. *New. Phytol.* **197**, 314–322 (2013).

838
839
840
841

842 **Acknowledgments**

843 We thank Prof. Doyle McKey, Prof. John Pannell, Dr. Clément Lafon-Placette and Dr.
844 Fabrice Roux for their comments on the manuscript, Y. Marie for the Enzymatic Methyl-
845 sequencing (EM-seq) conducted by the IGen Seq platform (ICM, Paris, France), C.
846 Chaparro (IHPE, Perpignan) for advice on the bioinformatic analysis of EM-seq data on
847 the IHPE Galaxy platform, C. Grunau (IHPE, Perpignan) for animating discussions in the
848 research network GDR 3E (Epigenetics in Ecology and Evolution) and the members and
849 coordinators (including S. Maury) of the COST action (European Cooperation in Science
850 and Technology) EPIgenetic mechanisms of Crop Adaptation To Climate cHange
851 (EPICATCH)—grant number CA19125 for discussions.

852

853 **Funding:** This project has received funding from the European Research Council (ERC)
854 under the European Union’s horizon 2020 research and innovation program (grant
855 agreement No ERC-CoG-2015-681484-ANGI) awarded to BP.

856

857 **Author contributions:** Each author’s contribution(s) to the paper should be listed (we
858 suggest following the CRediT model with each CRediT role given its own line. No
859 punctuation in the initials.

860

Conceptualization: BP, SMAury

861

Plant experiment and data collection: M.L., J.P., S.Marin

862

Bioinformatic analyses: N.L.-A

863

Data analyses: B.P., M.L., N.L.-A., P.M.

864

Graphical artwork: N.L.-A, B.P

865

Supervision: BP, SMAury

866

Writing—original draft: B.P., N.L.-A. and S.Maury

867

Writing—review & editing: all authors

868

869 **Competing interests:** All other authors declare they have no competing interests.

870

871 **Data and materials availability:**

872 All data are available in the main text or the supplementary materials. The EM-seq data
873 used in this study are available on the NCBI bioproject online repository. The names of
874 the repository/repositories and accession number(s) can be found below:
875 <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA773039>. Data and R protocols used for
876 the statistical analyses are available from the ZENODO Digital Repository

<https://doi.org/10.5281/zenodo.7603324>. Correspondence and requests for the seed material of the 123 epiRILs - see Table S5 in Colomé-Thatché et al. (25) for a list of accession names - should be addressed to the Institute Jean-Pierre Bourgin (IJPB - UMR1318 INRAE-AgroParisTech, Versailles, France, <https://publiclines.versailles.inrae.fr/catalogue/epiril>). No material transfer agreement is necessary for scientific work on accessions of the model plant *Arabidopsis*.

Figures and Tables

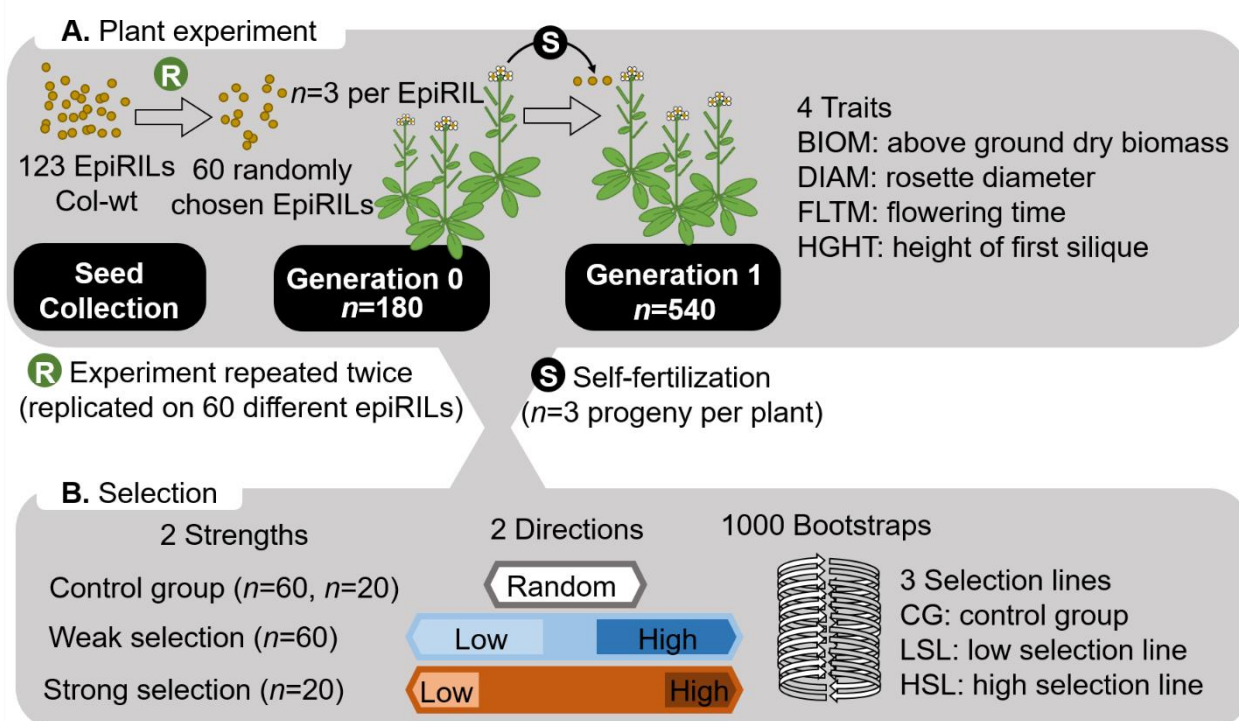
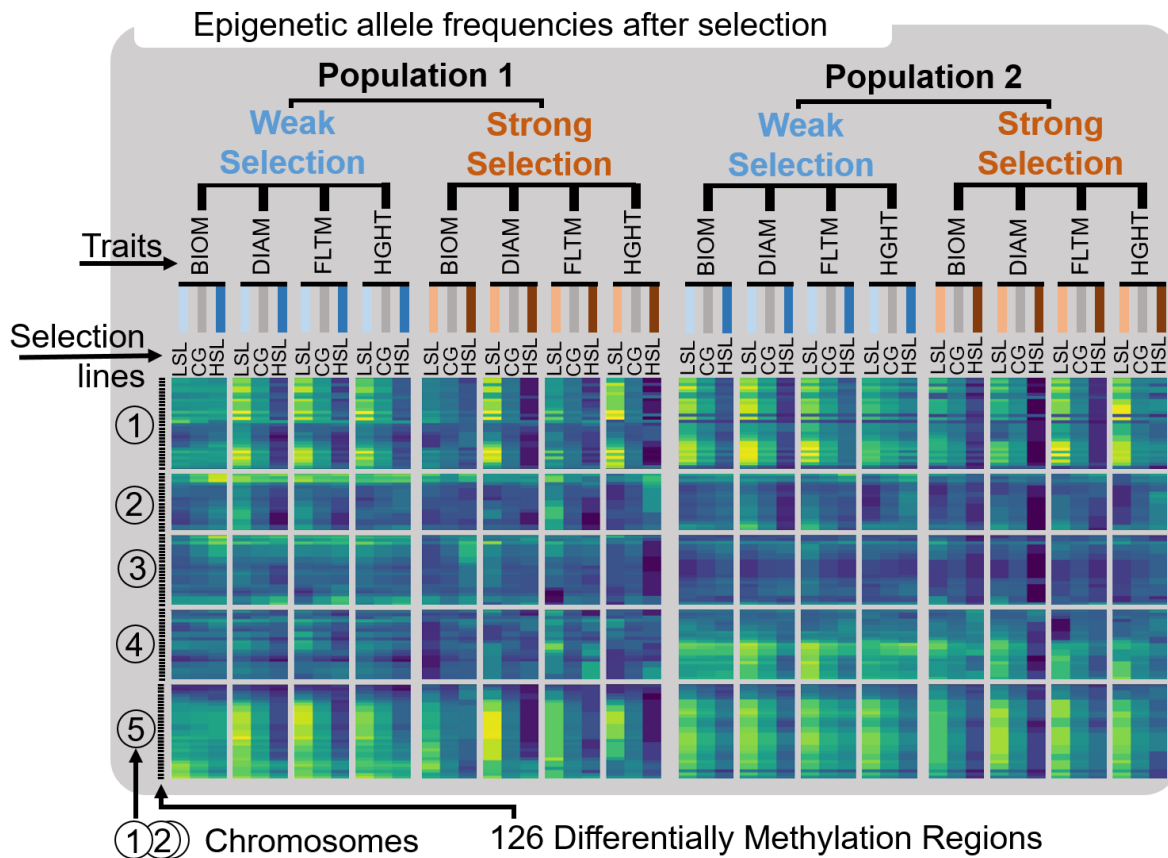


Fig. 1. Selection experiment drives phenotypic trait changes. (A), Plant experiment: Two sets of 60 different epiRILs (*Arabidopsis thaliana* Col-ddm1) were randomly assembled to build the two different experimental populations that were cultivated at different periods in the cultivation chamber. Seeds from the collection were grown to produce a buffer generation without selection and generate the seeds used to grow G0 plants. G1 plants – the progeny of G0 – were produced by self-fertilization. (B), Selection design: At G0, a control group was made of randomly selected plants. Plants with the highest and lowest trait values were respectively used to build the high and low selection lines. Weak and strong selection lines were respectively made up of 33% and 11% of the population. Corresponding progenies were used to build control groups and selection lines in G1, with differences between G1 lines reflecting responses to selection.



899
 900
 901 **Fig. 2. Selection experiment drives changes in epigenetic frequencies.** Changes in
 902 epigenetic allele frequencies after selection: The heat map presents DMR
 903 frequencies in G1 selection lines (LSL: low selection line and HSL: high selection
 904 line) and control groups (CG) for each of the four traits (BIOM: above ground dry
 905 biomass, DIAM: rosette diameter, FLTM: flowering time, HGHT: height of first
 906 silique) in populations 1 and 2. DMRs location on the five chromosomes are
 907 labelled on the left-hand vertical axis. Methylation frequencies of DMRs are
 908 represented by a colored band, with a gradient from blue (wild-type methylation
 909 status) to yellow (fixation of *ddm1-2* mutant methylation status). In each panel,
 910 DMR frequencies on a given chromosome are presented for the low, control and
 911 high selection lines. Each column corresponds to a trait. Frequencies were
 estimated on the basis of 1000 bootstraps.

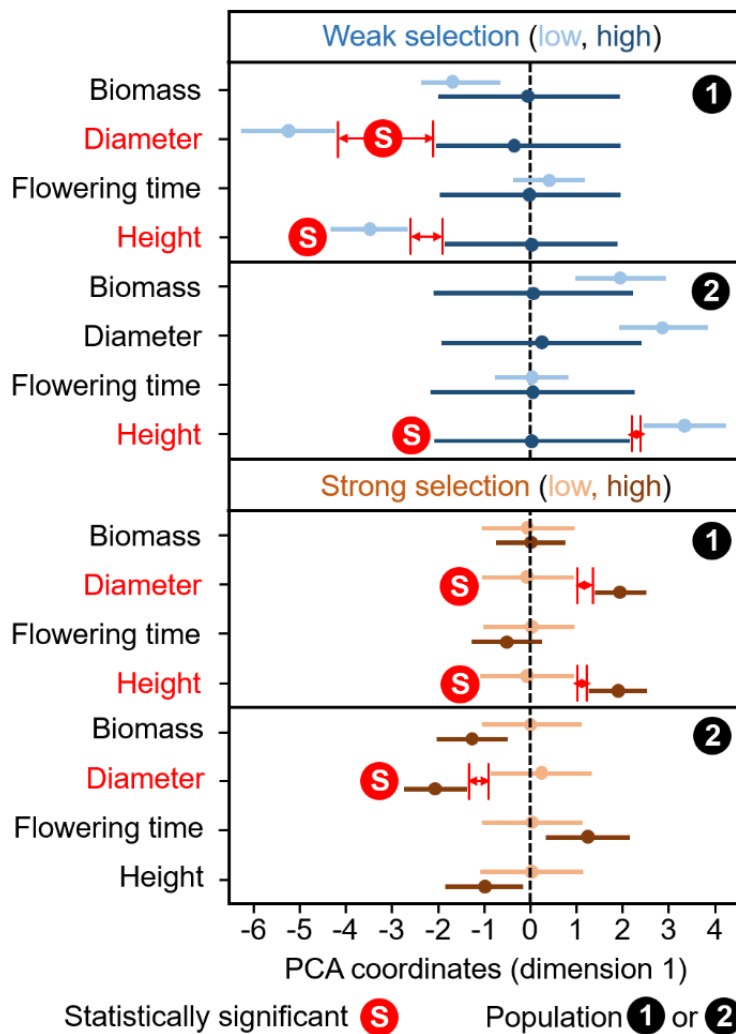


Fig. 3. Epigenetic divergence or its lack thereof between selection treatments.

Epigenetic divergence for each *Arabidopsis thaliana* epiRIL population (pop 1 and 2), each trait, each selection strength ($n = 60$ for each line under weak selection, $n = 20$ for each line under strong selection), and line (high and low selection lines) of plants. Principal Component Analysis coordinates for the first dimension (PC1), which synthesizes 18% and 22% of the global DMR variation in populations 1 and 2 respectively, are presented with their 95% confidence intervals. Divergence was considered significant following a conservative approach avoiding sampling bias: when 95% Confidence Intervals of PC1 values estimated on the basis of 1000 bootstraps did not overlap between upper and lower selection lines. Significant epigenetic divergent responses to selection are shown in red.

Table 1. Selection differential (S), response to selection (R) and heritability (h^2) for every selection line. Weak selection: selection by truncation of the upper and the lower third ($n = 60$) of *Arabidopsis thaliana* plants ranked by trait values to compose high and low selection lines (HSL and LSL, respectively). Strong selection: same method but selecting by truncation the upper and lower ca. 11% ($n = 20$) of plants. R is presented as trait change in trait units (g, cm, d) and percentage change. Narrow-sense heritability (h^2) calculated by the breeder's equation is in fact broad-sense heritability (H^2) in self-fertilizing plants. Selection was considered significant following a conservative approach avoiding sampling bias; when the 95% Confidence Intervals of average trait values estimated on the

935
936
937
938

basis of 1000 bootstraps did not overlap between upper and lower selection lines (Extended Data Table 3). Biomass: above-ground dry biomass (g). Diameter: rosette diameter (cm). Flowering time (d). Height: Height at first fruit (cm).

Pop	Trait	Weak selection						Strong selection					
		LSL			HSL			LSL			HSL		
		S	R	H ²	S	R	H ²	S	R	H ²	S	R	H ²
1	Biomass	0.07	NS	N/A	0.07	NS	N/A	0.16	NS	N/A	0.14	NS	N/A
	Diameter	0.87	0.42 (5%)	0.48	0.89	0.41 (5%)	0.46	1.90	0.59 (7%)	0.31	1.74	0.66 (8%)	0.38
	Flowering time	1.36	0.85 (3%)	0.63	1.34	0.80 (3%)	0.60	2.65	1.28 (5%)	0.48	3.41	1.81 (6%)	0.53
	Height	1.85	0.41 (4%)	0.22	1.85	0.41 (4%)	0.22	3.80	1.28 (11%)	0.34	3.94	1.50 (13%)	0.38
2	Biomass	0.11	NS	N/A	0.11	NS	N/A	0.24	NS	N/A	0.23	NS	N/A
	Diameter	0.97	0.55 (6%)	0.57	1.69	0.72 (8%)	0.42	1.76	0.67 (8%)	0.38	3.90	1.13 (13%)	0.29
	Flowering time	1.63	NS	N/A	1.72	NS	N/A	2.77	0.87 (3%)	0.31	4.88	0.62 (2%)	0.13
	Height	2.11	0.99 (8%)	0.47	1.91	0.90 (7%)	0.47	4.05	2.04 (16%)	0.50	4.60	1.47 (11%)	0.32

939