GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2011/05/19/genetics.111.128744.DC1

Genome-Wide Epigenetic Perturbation Jump-Starts Patterns of Heritable Variation Found in Nature

Fabrice Roux, Maria Colomé-Tatché, Cécile Edelist, René Wardenaar, Philippe Guerche, Frédéric Hospital, Vincent Colot, Ritsert C. Jansen, and Frank Johannes

DOI: 10.1534/genetics.111.128744

File S1

Supporting material:

Plant populations:

This study involved three different groups of plant material, i.e. the Col-wt and Col-*ddm1* parental lines of the epiRIL population, the Col-wt epiRILs and 10 natural accessions. Seeds for Col-wt and Col-*ddm1* parental lines were obtained as described in [1]. Construction of the Col-wt epiRILs has been fully described elsewhere [1]. Col-wt epiRILs correspond to the first subline of BC1-S5 (F7) plants. Due to few available seeds for some plant lines, 477 out of 505 Col-wt epiRILs were used in this study. To reduce maternal effects, seeds for the Col-wt and Col-*ddm1* parental lines and the Col-wt epiRILs were produced in the same greenhouse conditions at INRA Versailles.

In order to compare phenotypic diversity between epiRILs and a set of natural accessions, we first constructed a phenotypic space by running a Principal Component Analysis (PCA) on 20 quantitative traits scored on 240 worldwide natural accessions in greenhouse conditions (treatment without vernalization) [2]. Quantitative traits included morphological, phenological, architectural and fitness-related traits. Since quantitative traits were expressed in different units, PCA was run on a correlation matrix based on quantitative traits standardized to zero mean and unit variance (Systat 12 software). The phenotypic space was determined by the two first axes of the PCA explaining 47.6% of the phenotypic variation (first axis: 26.1%, second axis: 21.5%). For the purpose of this study, ten natural accessions have been chosen according to two conditions. Firstly, natural accessions should not require a vernalization treatment to induce flowering. Secondly, in order to maximize the phenotypic diversity observed at the worldwide scale, natural accessions should spread over the phenotypic space. This resulted in the selection of the following accessions: An-1, En-1, Gr-1, Hs-0, Is-0, Jm-0, Ka-0, Nd-1, Per-1 and Sap-0. To reduce maternal effects, seeds for the ten natural accessions were produced under controlled greenhouse conditions (16-h photoperiod, 20 degrees C) at the University of Lille 1.

Field experiment:

An experiment of over 6000 plants was set up according to a factorial randomized block design involving six blocks, where all combinations of two factors were included in each block. The first factor corresponds to the plant lines of each plant material group described above. The second factor corresponds to a density treatment simulating two levels of intra-specific competition, as previously described in [3]. The low-density (absence of competition) and high-density (presence of competition) treatments simulate two competitive environments frequently observed in natural populations of *A. thaliana*.

Each block was represented by 17 arrays of 66 individual wells (6 lines x 11 columns, $\oslash$4 cm, vol.$\approx 38 cm^3$) (TEKU, JP 3050/66) filled with damp standard culture soil (Huminsubstrat N3, Neuhaus). Twenty-two wells were left empty in the 17th tray, resulting in 1100 planting wells. Density treatment was randomly assigned to one-half of the planting wells. Each density treatment was an independent randomization of 10 replicates per parental line (n = 2), 1 replicate per Col-wt epiRIL (n = 477) and 3 replicates per natural accession (n = 10).

For both low-density and high-density treatments, three seeds were sown in the central position of each well. In the high-density treatment, each central focal plant were surrounded by six *A. thaliana* neighbors evenly spaced 1cm away in each direction and all plants within a well belonged to the same plant line. Seeds were sown a block per day between from 7-12 March 2007 in a greenhouse mimicking the outdoor conditions (no additional light or heating) and protecting seeds from rainfall. Germination date was monitored in each well 4th and 8th day after sowing. *A. thaliana* seeds in the central position that had not germinated 14 days after sowing (proportion = 0.9%) were replaced by extra-seedlings of the same plant line from the same block. *A. thaliana* seeds at the surrounding positions in the high-density treatment that had not germinated 14 days after sowing (proportion = 3.1%) were replaced by extra-seedlings of the same plant line from the same block or from other blocks. Central focal seedlings were thinned to one per well 15 days after the sowing, keeping the first germinated seedling.

During all the growing period in the greenhouse, arrays were rotated every other day to minimize potential effects of micro-environmental variation. Plants were transplanted a block per day between from 2-7 April 2007, i.e. 26 days after sowing, into a tilled common garden located at the University of Lille 1. The six blocks were arranged at 75cm spacing in the common garden. Each block was represented by grid of 11 lines by 100 columns (Supplementary Figure 2a), with central focal plants spaced at 10cm to avoid competition among focal plants (Supplementary Figure 2b). The plants were watered for a week to ease the acclimatizing to common garden conditions. Vertebrate herbivores were excluded by two successive fences. Molluscicide (PhytorexJ, Bayer Jardin) was scattered across the common garden to prevent slug attacks.

Phenotyping:

We measured several quantitative traits that have been described as non-collinear and as related to adaptation in *A. thaliana* [2]. In the common garden, central focal plants were monitored for floral transition every day from April 14 2007 to 28 May 2007. Flowering time was scored as the number of days between germination and the appearance of the first open flower. The experiment stopped when all plants senesced, i.e. all fruits (i.e. pod) were mature. At the end of the experiment, the aboveground portion of each focal plant was collected and stored at 4 degrees C for further phenotyping.

Since one block located at the edge of the common garden was invaded by *Trifolium campestre* in June, we were unable to collect the focal plants in this specific block. In the remaining five blocks, three architectural traits were measured on all focal plants: height from soil to first fruit, height of the main stem, maximal plant height. We note here that, in some epiRILs, we observed more resources allocation to primary branches than to the main stem. As a consequence, maximal height appeared to correlate poorly with the height of the main stem. For this reason, we chose to measure both traits seperately. Dry above-ground biomass expressed in mg was measured on all focal plants. In three randomly chosen blocks, four fitness-related traits were measured

for all focal plants in the high-density treatment. We separately counted the number of fruits produced on the main stem and the primary branches on the main stem. Since the length of a fruit strongly correlates with the number of seeds contained within it [4], we also estimated separately the fruit size (calculated as the average of three randomly chosen fruits) on the main stem and the primary branches on the main stem.

Heritability estimates (H^2):

Heritability estimates in this study focused on main effects by pooling plants from both density treatments. Hence, let $y_{i,j}$ denote the treatment-adjusted (i.e. competition and block effect) phenotypic value for the jth plant of the ith line. Phenotypes were log2-transformed when deemed appropriate. We modeled the line-effect using a random intercepts model:

$$y_{i,j} = \beta_0 + b_i z_{i,j} + \epsilon_{i,j}, \tag{1}$$

where β_0 is a common fixed intercept, b_i is the random intercept of the ith line, $z_{i,j}$ is an index variable and $\epsilon_{i,j}$ is the error. Assume that $b_i \sim N(0, \Psi^2)$, $\epsilon_{i,j} \sim N(0, \sigma_i^2)$, and $\mathrm{cov}(\epsilon_{i,j}, \epsilon_{i,j'}) = 0$. Estimates were obtained by maximum likelihood ($lme4$ library in R, [5]). We evaluated the line-effect by testing $H_0 : \Psi^2 = 0$ vs. $H_A : \Psi^2 > 0$ via the likelihood ratio test. Broad-sense heritability was calculated as

$$\widehat{H^2} = \widehat{\Psi^2}/\widehat{\sigma^2}(\vec{y}). \tag{2}$$

Standard errors (SE) for $\widehat{H^2}$ were obtained using 3000 bootstrap samples ($mcmcsamp$ function in R).

Estimates of the number of QTL:

Estimates of the number of QTL (N) were obtained using [6]:

$$N = \frac{3D^2}{(1-2\tau)^2} \frac{1 + (1-2\tau)^2(2\bar{r}-1)/(2\bar{r}+1)}{16\widehat{\Psi^2} + 3D^2(2\bar{r}-1)/(2\bar{r}+1)}, \tag{3}$$

where D is the parental mean difference, $\bar{r} = 0.44$ is the average recombination fraction and τ is the average transgression potential between the parents. In this approximation we assume full stability of induced

(epi)alleles. This assumption could be relaxed to account for possible reversion of epialleles or other time-dependent behaviors [6]. The standard errors (SE) were obtained using a nonparametric bootstrap approach.

Average estimates of the number of QTL and H^2:

For Fig. 1C we calculated the average number of QTL and heritability for two different trait categories. Let $\hat{E}_j$ denote the estimate (either for heritability or QTL number) for the j^{th} of n traits. The average is

$$\overline{\hat{E}} = \sum_{j=1}^{n} w_j \hat{E}_j, \tag{4}$$

where w_j is a sample size weight $w_j = N_j / \sum_j N_j$. To obtain the standard errors (SE) of $\overline{\hat{E}}$, we calculate $\sigma(\overline{\hat{E}})^2$. We approximate this quantity using

$$\sigma(\overline{\hat{E}})^2 = \sum_{j=1}^{n} w_j^2 \sigma(\hat{E}_j)^2 + 2 \sum_{i<j} w_j w_i \mathrm{cov}(\hat{E}_j, \hat{E}_i). \tag{5}$$

Since the genetic correlations between traits are only modest, the terms $\mathrm{cov}(\hat{E}_j, \hat{E}_i)$ are negligible. The variance terms $\sigma(\hat{E}_j)^2$ $(j = 1, ..., n)$ are obtained from 3000 stratified bootstrap samples.

References

[1] Johannes F. *et al.*, *PLoS Genetics*, **5**, e1000530 (2009).

[2] Reboud X., Le Corre V., Scarcelli N., Roux F., David J.L., *et al.*, NRC Research Press, Ottawa, Ontario, pp 135-142 (2004).

[3] Weinig C., Johnston J., German Z.M., Demink L.M., *The American Naturalist*, **167**: 826-836 (2006).

[4] Roux, F., J. Gasquez and X. Reboud, *Genetics* **166**: 449-460; doi:10.1534/genetics.166.1.449 (2004).

[5] Pinheiro J.C. and D.M. Bates, Mixed-Effects Models in S and S-PLUS, Springer-Verlag (2000).

[6] Johannes F. and M. Colomé-Tatché, *Genetics*, DOI:10.1534/genetics.111.127118 (2011).

File S2

Supporting Data

File S2 is available as a compressed folder at http://www.genetics.org/content/suppl/2011/05/19/genetics.111.128744.DC1.

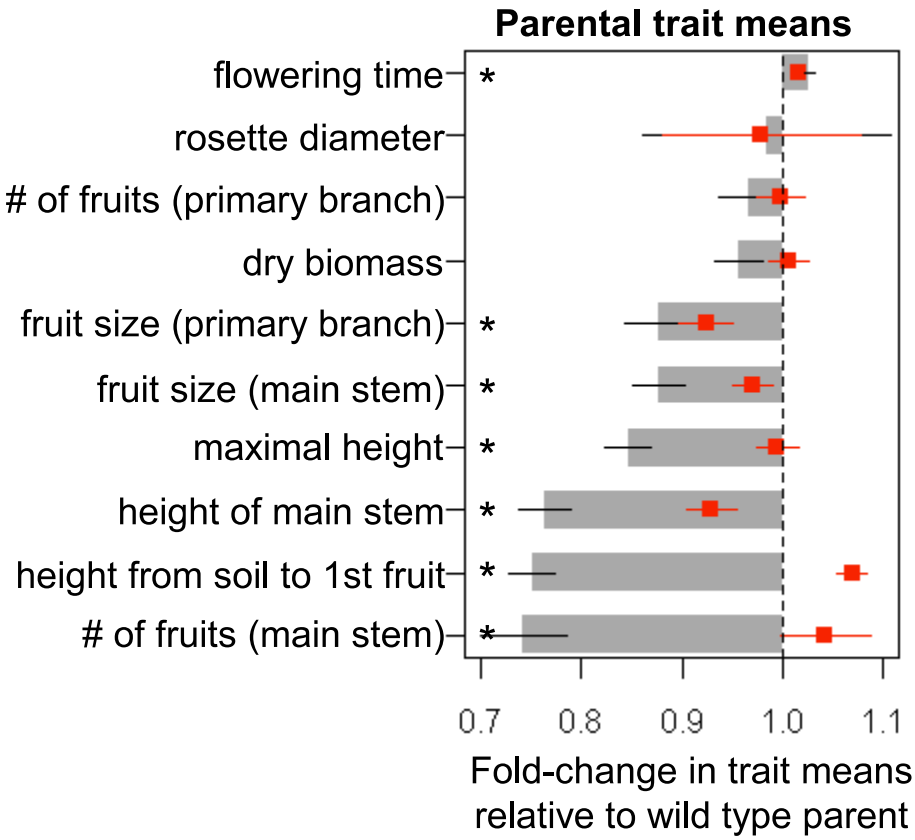

Figure S1 Fold change (±SE) in trait means of mutant parent (gray bars) relative to the wt parent (dashed line); asterisks indicate significant two-sided t-test at $P < 0.05$ testing for differences between the parents. For comparison, we also plot the fold-change (±SE) of the epiRIL trait means relative to the wt parent (red squares). The epiRIL means are typically closer to those of the wt.

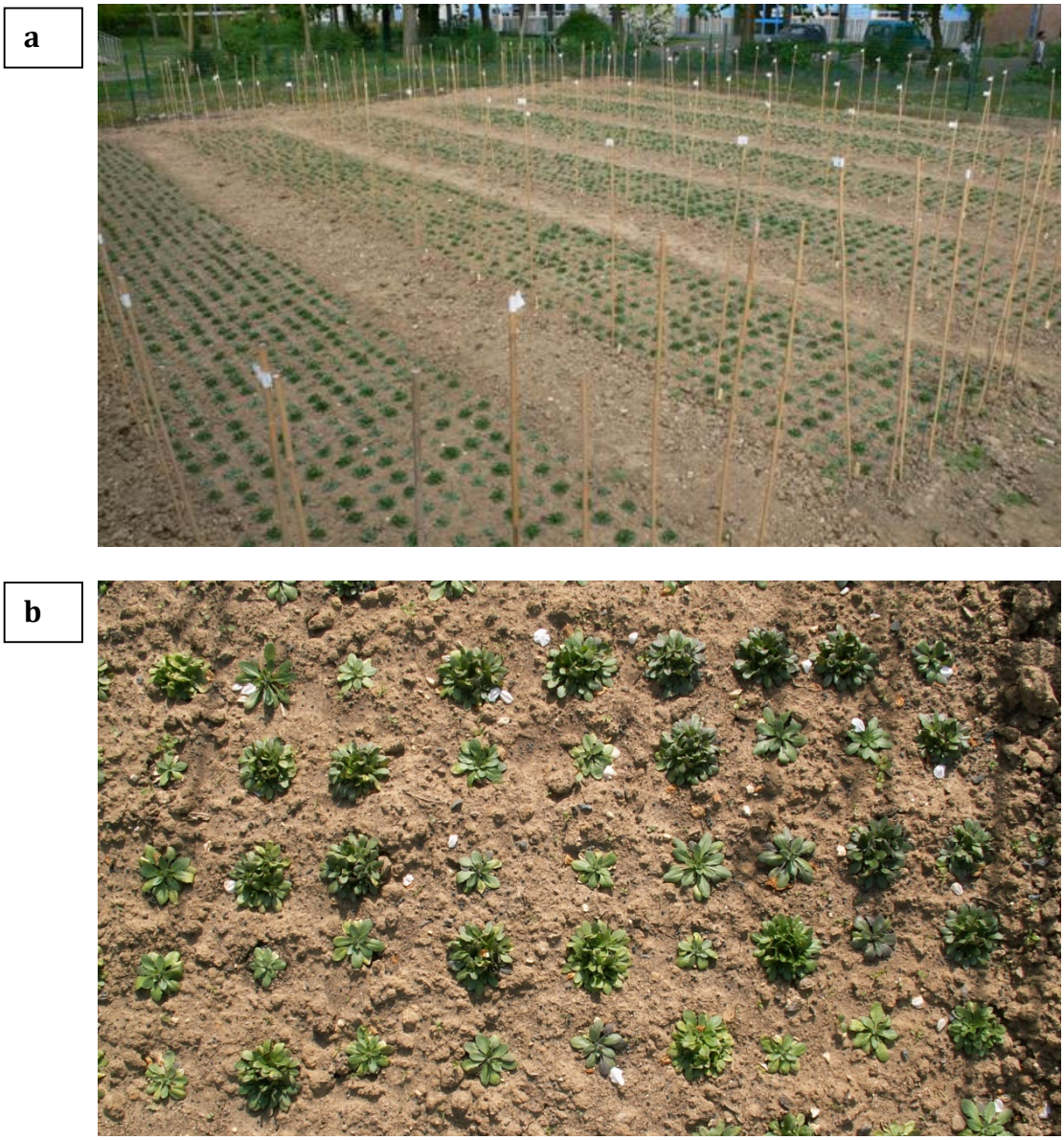

Figure S2 Field experiment. a. Six experimental blocks, each being represented by grid of 11 lines by 110 columns. b. 10-cm spacing among focal plants.

Table S1 Parental phenotypic means

trait name	Col-*ddm1*		Col-wt		*t* value	df	*p* value
	n	mean	*n*	mean			
flowering time	117	43.40	116	42.34	-4.00	223.17	< 0.0001
rosette diameter	59	4.90	58	5.00	1.15	110.75	0.20
# of fruits (primary branch)	28	7.33	30	7.59	1.07	51.52	0.29
dry biomass	97	5.73	96	5.98	1.77	186.30	0.08
fruit size (primary branch)	28	10.44	30	11.90	3.24	54.03	< 0.0001
fruit size (main stem)	28	12.30	30	14.03	4.48	55.93	< 0.0001
maximal height	95	27.75	95	32.77	6.29	164.94	< 0.0001
height of main stem	96	27.06	95	35.46	7.78	152.76	< 0.0001
height from soil to 1st fruit	95	7.11	95	9.47	9.81	158.78	< 0.0001
# of fruits (main stem)	28	12.30	30	14.03	4.91	51.82	< 0.0001

Provided are the parental phenotypic means along with the results of a two-sided t-test evaluating significant differences between the Col-*ddm1* and Col-wt parents. In the case of unequal trait variances, a Welch modification to the degrees of freedom (df) is used.

TABLE S2 Heritability estimates in the epiRILs and natural accessions

trait name	Population	*n*	H^2	(s.e)	X^2	df	*P* value
flowering time	epiRILs	5567	0.2997	0.0109	1225.14	1	< 0.0001
	accessions	356	0.8602	0.0142	610.91	1	< 0.0001
fruit size (main stem)	epiRILs	1317	< 0.0001	0.0054	0	1	0.99804
	accessions	86	0.6017	0.0790	48.54	1	< 0.0001
height from soil to 1st fruit	epiRILs	4535	0.1571	0.0112	408.38	1	< 0.0001
	accessions	292	0.4164	0.0430	118.42	1	< 0.0001
maximal height	epiRILs	4543	0.0628	0.0102	138.96	1	< 0.0001
	accessions	293	0.4032	0.0428	113.56	1	< 0.0001
height of main stem	epiRILs	4561	0.0393	0.0091	87	1	< 0.0001
	accessions	292	0.3593	0.0447	95.89	1	< 0.0001
fruit size (primary branch)	epiRILs	1326	0.0392	0.0060	16.49	1	< 0.0001
	accessions	86 0.	1858	0.0861	9.15	1	0.00249
# of fruits (main stem)	epiRILs	1330	< 0.0001	0.0050	0	1	0.99633
	accessions	88	0.1394	0.0807	6.52	1	0.01067
dry biomass	epiRILs	4586	0.0184	0.0083	39.18	1	< 0.0001
	accessions	294	0.0075	0.0229	1.09	1	0.29624
rosette diameter	epiRILs	2789	0.0160	0.0116	*	*	*
	accessions	177	0.0069	0.0378	*	*	*
# of fruits (primary branch)	epiRILs	1333	< 0.0001	0.0030	0	1	0.99817
	accessions	87	< 0.0001	0.0262	0	1	0.99991

* No values available because H^2 were based on average estimates in this case (see Supporting Methods). However, 95% confidence (±1.96 x s.e.) intervals include zero suggesting non-significance.

Provided are the broad-sense heritability estimates (H^2) for the epiRILs and natural accessions. Estimates were obtained from a random effects model as described in the Supporting Methods. * No values available because H^2 were based on average estimates in this case (see Supporting methods). However, 95% confidence (±1.96 x s.e.) intervals include zero suggesting non-significance.