Weak evidence for heritable changes in response to selection by aphids in *Arabidopsis* accessions

Marc W. Schmid[1,*], Klara Kropivšek[2,3,†,*] , Samuel E. Wuest[2,3,4,††], Bernhard Schmid[3,4,5], Ueli Grossniklaus[2,3]

[1] MWSchmid GmbH, 8750 Glarus, Switzerland. [2] Department of Plant and Microbial Biology, University of Zurich, 8008 Zurich, Switzerland. [3] Zurich-Basel Plant Science Center, University of Zurich, ETH Zurich, University of Basel, 8092 Zurich, Switzerland. [4] Department of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland. [5] Department of Geography, University of Zurich, 8057 Zurich, Switzerland.

* These authors contributed equally to this work.

[†] Present address: Laboratory for Environmental and Life Sciences, University of Nova Gorica, 5000 Nova Gorica, Slovenia

[††] Present address: Agroscope, 8820 Wädenswil, Switzerland.

Correspondence and requests for materials should be addressed to U.G. (email: grossnik@botinst.uzh.ch).

Key words

Genetics, epigenetics, selection, evolution, plants, herbivory, aphids, *Arabidopsis*, DNA methylation

Abstract

In plants, transgenerational inheritance of certain epialleles has been observed but experimental evidence for selection of epigenetic variation independent of genetic variation is scarce. We extended an experiment simulating selection in response to herbivory in *Arabidopsis thaliana* to assess a potential contribution of epigenetic variation to the selected phenotypes within three accessions. To minimize maternal effects, we grew offspring from replicate populations and their ancestors for two generations in a common environment and assessed the phenotypes in the second generation. We found weak evidence for the selection of epigenetic variation: bolting time differed significantly in one accession. Significant differences between maternal lines suggested random residual or novel genetic and/or epigenetic variation. Our results are in conflict with those of a recent study reporting that environment-induced heritable variation is common in *Arabidopsis*. Reanalyzing the data from that study showed that the reported findings resulted from a mix-up of accessions and, thus, reflected genetic rather than epigenetic variation between accessions. To avoid future misinterpretations of

studies investigating epigenetic inheritance, we provide guidelines to design experiments that clearly differentiate between epigenetic and genetic variation and distinguish standing variation from *de novo* variation acquired during an experiment.

Introduction

Modifications of DNA and histones represent epigenetic marks associated with active or repressed gene expression and play an important role in plant development and responsiveness to the environment. Unlike in animals, plant germlines are not set aside early in development but form later, when somatic cells get committed to form gametes (1). As a result, epigenetic marks affected by environmental conditions can potentially be inherited. Therfore, epigenetic marks may contribute to adaptive responses in changing environments. Furthermore, spontaneous changes in DNA-methylation, a mark associated with some naturally arising heritable epialleles in the plant model *Arabidopsis thaliana* (e.g., 2, 3, 4), occur much more frequently than genetic mutations (5, 6, 7). This suggests that the rates of epimutations might be sufficiently high, differentiating epigenetic from genetic variation, yet low enough to maintain a heritable response to selection.

The potential role of epigenetic variation in adaptation depends on several factors: the rate of spontaneous epimutations, the influence of environmental effects on epigenetic variation, the heritability of epimutations, and the impact of epigenetic variation on the phenotype. Importantly, to contribute to adaptation, epigenetic variation must be sufficiently stable, i.e., heritable over several generations, so it can be subject to selection. In non-model species, DNA methylation patterns assessed with reduced representation sequencing methods were frequently found to be associated with distinct environments or treatments. However, epigenetic variation was generally correlated with genetic variation (8, 9, 10, 11, 12). Likewise, high-resolution genome-wide studies in *Arabidopsis* have revealed extensive variation in DNA methylation patterns among different populations, but this variation was mostly associated with underlying genetic differences (13, 14). Nonetheless, studies on epigenetic recombinant inbred lines (epiRILs) suggest that epialleles can significantly contribute to phenotypic variation in plants (15, 16). Thus, while there are examples of pure epialleles (17) whose frequency is not related to genetic variation, the role of spontaneous epigenetic variation in adaptation remains controversial (18, 19, 20, 6). This is largely because, in most species, it is difficult to distinguish between epigenetic variation that results from genetic differences and pure epigenetic variation, which is independent of the genotype (21). Disentangling genetic and epigenetic effects bears most promise in clonally reproducing species, such as dandelions or strawberries (22, 23), or highly inbred species, such as *Arabidopsis*.

In a previous experiment with *Arabidopsis* (4), we simulated selection in a rapidly changing environment. We compared phenotypic traits and epigenetic variation between populations grown for five generations under selection with their genetically identical ancestors in a common environment. Selected populations of two distinct genotypes exhibited significantly different flowering times and plant architectures when compared with their ancestral populations, and these differences persisted for at least three

generations in a common environment. In parallel, we observed a reduction in both, epigenetic diversity and changes in DNA methylation patterns, some of which were associated with the observed phenotypic changes. These findings indicated that epigenetic variation was subject to selection. The genotypes used in the described experiment were recombinant inbred lines (RILs), formed by hybridizing two different accessions, followed by several generations of inbreeding. Therefore, most of the epigenetic variation in these RILs was likely retained as standing variation from the original hybridization event (24) and did not accumulate during the selection experiment through spontaneously occurring or environmentally induced epimutations (4).

Here, to assess the potential for epigenetic differentiation between selection treatments, we used a similar approach using *Arabidopsis* accessions from a previous selection experiment that investigated the effects of aphid herbivores on plant defensive traits (25). In this selection experiment, mixtures of genetically uniform *Arabidopsis* accessions were grown for five generations in either insect-free cages (control) or cages inoculated with different aphid species or a mixture of these species. After five generations, genetic variation was significantly reduced, and populations were dominated by three accessions to various degrees (25). To examine whether epigenetic variation was also selected within these three accessions (Ma-0, Wei-0, Wil-3), we compared phenotypic traits of selected populations (i) to the original founder population, (ii) to the control population, and (iii) between two single-species aphid treatments. From each selection treatment, we used six independent replicate populations after five generations of selection and grew them together for two generations in a randomized planting pattern in a common environment. To minimize confounding maternal effects, we assessed the phenotypes only in the second generation. We further compared genome-wide DNA methylation states in a subset of mother plants used to set up the common-environment experiment.

We designed the experiment to detect effect sizes of 10% or more at the significance level of $P < 0.05$ with a statistical power of $beta > 0.80$. While we found highly significant differences between accessions, replicate populations, and across five experimental blocks (with plants generally growing worse in two blocks compared to the other three), we observed only few significant differences between selection regimes (control and the two different aphid treatments). These differences were weak, with the percentage sum of squares explained by the contrasts being less than 1% and P values ranging between 0.05 and 0.01. We found minimal differences in DNA methylation patterns within accessions between plants from different selection regimes, with 0.009% of all tested cytosines showing significant differences after correction for multiple testing (FDR < 0.05). In contrast to our previous study (4), there was no reduction in the amount of epigenetic variation in response to selection. In summary, we found evidence that some within-accession differences caused by the selection regimes persisted for two generations in the common environment. However, the extent of these differences was low. Thus, we conclude that the evidence for selection of epigenetic variation in this experiment is weak.

In strong contrast to our findings, a recent study claimed that environment-induced heritable variation is common in *Arabidopsis* (26). We were surprised about this finding because, in our experiment, the aphid treatments had a strong impact on plant fitness and survival but not on epigenetic differentiation. Thus, if environment-induced heritable

variation was indeed so common in this model species, we would have expected to observe it, too. We noticed that the design of the study by Lin and colleagues (26) was elegant but susceptible to potential errors that confounded treatments with genetic variation at the beginning of the experiment. We used their RNA-seq data to assess whether the analyzed plants displayed any evidence of residual genetic variation that was potentially confounded with the treatments. Our analysis showed that about 10% of the plants that we could assess belonged to accessions differing from the ones they had been assigned to. This error did not occur at random but stemmed from a mix-up of two selection treatments. Thus, the large effects assigned to environment-induced heritable variation were very likely due to genetic differences between accessions.

In summary, while epialleles can be subject to selection and contribute to adaptation (4), the frequency of these events remains unclear, and epigenetic differentiation in response to selection seems to require standing epigenetic variation within accessions. In contrast, evidence for a role of environment-induced epigenetic changes in adaptation in the model plant *Arabidopsis* remains weak.

Results

Impact of Selection Treatments on Phenotypic Traits

As described above, mixtures of genetically uniform *Arabidopsis* accessions were grown for five generations, either under control conditions or inoculated with different aphid species (25). Thereafter, the populations were dominated by the three accessions Ma-0, Wei-0, and Wil-3. Phenotypic differences within uniform genetic backgrounds that were subjected to different treatments would indicate the selection of epigenetic variation. Thus, to investigate whether the treatments led to phenotypic differences within the three dominant accessions, we tested whether the selection treatments had an impact on the time to bolting, the rosette diameter at bolting, the number of rosette leaves at bolting, the leaf blade area, and the trichome density (Tables 1-2, Figures 1-3). Besides the selection treatment and its contrasts, the model also included the experimental blocks (n = 5), the accessions (n = 3), the populations (n = 18+3 ancestrals), and maternal lines ("mother", n = 261), along with interactions of the selection treatment contrasts with accessions. Additionally, we conducted analyses without the Wil-3 accession because this accession had low replication at the level of maternal lines. We also ran separate analyses for each accession. Additional models and power estimations can be found in the Supplementary Information (Supplementary data S1 and Supplementary data S2).

The term "block", especially the comparison between blocks 1–3 and blocks 4–5, was highly significant in all analyses and generally explained most of the total phenotypic variation (up to nearly 50% for leaf blade area, Table 2). The time to bolting was an exception because here, the term "accession" explained most of the variation (23.2%), followed by "population" (8.8%), "mother" (8.1%), and "block" (6.5%). The terms "accession" and "population" were also consistently significant and explained a large fraction of phenotypic variation. Effects of "population" were likely due to differences in

composition because accessions were not balanced between populations. Accordingly, the interaction between "accession" and "population", which would indicate differences between populations within accessions, was rarely significant and explained less than 1% of all variance. Additionally, the term "mother" was occasionally significant, accounting for 5% to 10% of the total phenotypic variance (Tables 1–2). Compared with these previous terms indicating environmental and genetic influences, the selection treatment and its contrasts – as well as their interactions with the previous terms – were rarely significant and explained less than 1% of the total phenotypic variation (Table1–2, Figures 2–3). Furthermore, the significant differences in rosette diameter between the control and aphid treatments and between the two aphid treatments (Table 1) disappeared when we excluded the accession Wil-3 (with low replication at the level of maternal lines) from the analysis (Table 2). In this case, only the interaction between the two aphid treatments and the two accessions Ma-0 and Wei-0 remained significant (Table 2). Separate analyses for each accession revealed significant differences in bolting time between the control and the aphid treatments in Ma-0 (Figures 2–3). We also observed a significant difference in trichome densities between the two aphid treatments in accession Wei-0; however, these were driven by some extreme values (Figure 3). We found several significant differences between selection treatments in accession Wil-3 (Figures 2–3) but, again, considering the low level of replication at the level of maternal lines in this accession, we interpret these differences with caution.

Limited Impact of Aphid Treatments on DNA Methylation Patterns and Epigenetic Variation

To more directly assess whether epigenetic variation could be responsible for the phenotypic differences observed between ancestral and selection treatments, we characterized changes in the genome-wide methylomes, taking DNA methylation as a proxy for epigenetic variation. We investigated the methylome of a subset of Wei-0 maternal lines from generation G6 (n = 34), providing a more direct readout of the aphid treatment than would be obtained with plants from generation G7. We assessed the DNA methylation patterns of these plants by whole-genome bisulfite sequencing. In total, we tested 3,573,936 cytosines for differential methylation between ancestrals and the two aphid-selection treatments. The number of significantly differentially methylated cytosines was similar in all comparisons: it ranged from 101, when the two aphid treatments were compared, to 131, when the *L. erysimi* treatment was compared with the ancestrals. Overall, about 0.009% of all tested cytosines showed significant differences, and they were dispersed across the entire genome (Figure 4A). We also tested whether the variation of DNA methylation, that is the extent of epigenetic variation, was reduced by the selection treatments, an effect of selection that we had previously observed in a similar experiment (4). However, we did not find any evidence for a reduction nor an expansion of variation in DNA methylation between ancestrals and the two aphid-selection treatments in any of the three contexts of DNA methylation: CG, CHG, and CHH (Figure 4B).

To understand the potential effects of differential cytosine methylation, we characterized all 308 differentially methylated cytosines (DMCs) with respect to sequence context and position relative to genes. The majority of DMCs was in the CG context (87.6%), and

only few were found in the CHG and CHH contexts (6.2% each). Considering all cytosines of the *Arabidopsis* genome, about 13%, 15%, and 72% are found in the CG, CHG, and CHH contexts, respectively (4). Thus, there was a clear enrichment of DMCs in the CG context, similar to our finding in the previous selection experiment (4) and consistent with the fact that methylation in the CG context shows the highest variability (27). Almost all DMCs were located within or less than 2 kb away from genes (87.3%) or genes and transposable elements (98.1%). Together with the clear enrichment of the CG-context, this suggests that most DMCs were related to differences in gene body methylation (gbM; 28, 29). We tested whether genes with certain functions were enriched for DMCs using gene ontology (GO) enrichment analysis (Table 3). There was an enrichment of several terms related to epigenetic regulation: "histone methylation", "DNA methylation", and "chromatin silencing". The remaining terms could be associated with transport and growth in general, for example "positive regulation of cell proliferation", "secondary cell wall biogenesis", and "organic substance transport". In conclusion, although we identified only few DMCs, they were preferentially found in genes associated with epigenetic regulation and transport and growth processes. However, the overall limited differentiation and the absence of reduced variation in DNA methylation suggest that there was only weak selection of specific DNA methylation sites.

The Large Effects Observed by Lin and Colleagues Are Not Due to Environment-Induced Heritable Variation but Due to Differences in Genotype

Compared with our study, which provided weak evidence for the selection of epigenetic variation, a recent report claimed that environment-induced heritable variation is common in *Arabidopsis* (26). We were surprised by this difference in the outcomes the two experiments, which both aimed at investigating the potential role of epigenetic variation in adaptation in this model species. Our aphid treatments strongly impacted plant fitness (25), indicating that an induction of heritable epigenetic variation within genetically uniform *Arabidopsis* populations, as suggested by Lin and colleagues (26), was possible.

We wondered if issues with the experimental design may have led to the strong effects in the study of Lin and colleagues (26). Using their RNA-seq raw data of 190 out of a total of 4,032 plants, we searched for residual genetic variation within accessions. Their RNA-seq data supposedly included four accessions. We compared pairwise genetic distances and identified the accessions most closely related to the experimental plants. We used 10,000 randomly drawn genes for the analyses and were left with up to 13,418 variants after filtering for coverage between 10–1,000 in at least four plants per treatment and accession. Visualizing pairwise genetic distances between all plants revealed that two presumed accessions, Abd-0 and TRE-1, were each represented by two genetically distinct clusters (Figure 5A). When we visualized each accession separately, highlighting the treatments, we found that the genetically distinct groups had been associated with different selection treatments, leading to an unwanted confounding of selection treatments with accessions. Specifically, in the high Cd treatment, the plants labeled as

Abd-0 were in fact from another accession and the same happened in the drought treatment where plants labeled as TRE-1 came from another accession (Figure 5B). When we compared the genomes of these experimental plants to those of publicly available accessions, we found that the genetically distinct plants from the high Cd treatment belonged to the accession Ws-2 and those from the drought treatment belonged to the accession Fei-0, both accessions that were otherwise included in the experiment (Supplementary Data S3). Notably, when we assigned the experimental plants to publicly available accessions, the distances to the closest accessions were between 0.0003 and 0.0034 (0.0019 on average), compared with distances to accessions Abd-0 and TRE-1, respectively, ranging from 0.358 to 0.451 (0.414 on average).

To confirm our finding of a potential mix-up leading to treatment–accession confounding, we used either the original or the newly assigned accession labels to split the pairwise distances between experimental plants into three classes: (i) within lineage (i.e., (grand)-parent-to-offspring), (ii) within accessions, and (iii) between accessions. The former two should both yield values close to zero as plants within lineages and within accessions should be genetically (nearly) identical to each other, whereas distances between accessions can be substantial (Figure 5C). These expectations were not met with the original accession labels; instead, we found large distances in all three classes. In contrast, the same evaluation with the re-assigned accession labels Ws-2 (for Abd-0 in high Cd treatment) and Fei-0 (for TRE-1 in drought treatment) confirmed the expectations of close to zero distances in classes (i) and (ii), and large distances in class (iii). Thus, it appears that Lin and colleagues (26) inadvertently had mixed up accessions when assigning plants to the two mentioned treatments.

In summary, our analysis reassigned 20 out of 190 plants (~ 10.5%) to new accessions. This error did not occur at random but stemmed from the wrong treatment assignment of plants from two accessions. This misassignment explains the large effects interpreted as environment-induced heritable variation by Lin and colleagues (26) that, as we show here, are really effects of genetic differences between accessions.

Discussion

In this study, we investigated whether selection by different aphid species within genetically uniform *Arabidopsis* accessions can lead to heritable phenotypic variation that is stable in the absence of selection for at least two generations, indicative of adaptation. Such (meta)stable phenotypic variation is expected to be predominantly epigenetic in nature, defined as a mitotically and/or meiotically heritable change in gene expression without a change in DNA sequence (30). As *Arabidopsis* is highly inbreeding, genetic variation within an accession is expected to be rare, while epigenetic differences, e.g., in DNA methylation, are more common (31, 32). We investigated epigenetic variation at multiple levels: differences between selection treatments within accessions would indicate epigenetic differentiation, while differences between replicate populations of these accessions could arise from either standing (epi)genetic variation at the beginning of the selection experiment or from *de novo* gained (epi)genetic variation developed

during the experiment, with genetic variants expected to be extremely rare in the case of *Arabidopsis*.

In our analyses, we found only weak evidence for epigenetic differentiation due to selection treatments, expected to manifest as significant phenotypic differences in the second generation in a common environment or as clear differences in the methylomes analyzed in our experiment. Except for one case, the differences were either limited to one accession with low replication at the level of mother plants or depended on the presence of a few extreme measurements. While differences between experimental blocks often explained the largest amount of the observed variation in phenotypes, this was different in the case of "time to bolting". This finding is in line with the notion that flowering time in *Arabidopsis* has a strong genetic component (33). Similarly, we consistently observed strong differences between the accessions, again emphasizing the genetic component of trait variability in this model species.

While we found little evidence for differences between populations within accessions, we found significant differences between maternal lines. In genetically uniform accessions, there are three possible explanations for these maternal line effects: (i) maternal effects, (ii) initial genetic or epigenetic differences between maternal lines that resulted in founder effects, and/or (iii) a differential (but selection treatment-independent) gain of genetic or epigenetic variation between maternal lines during the five generations of the selection experiment.

Initial genetic (or epigenetic) differences between maternal lines (or populations) may have two reasons. First, seeds of single accessions obtained from NASC may not have been derived from homozygous, genetically uniform plants. Second, there could have been genotyping errors, occurring during the identification of plants representing the accessions of interest. The latter case applied to the study of Lin and colleagues (26), in which two presumed accessions, Abd-0 and TRE-1, were thought to show strong evidence of environment-induced heritable variation while the other two, Ang-0 and Tol-0, did not. Plants of the two accessions showing no effect were genetically very similar to each other and were found identical to their accessions using publicly available data from over 1,000 accessions (1001genomes.org). However, plants from the presumed accessions with strong effects, Abd-0 and TRE-1, showed strong genetic differences between treatments, which results from a mix-up of accessions while preparing or conducting the experiment.

While experimental errors can only be addressed by verifying the material's identity at each step (which can be costly and challenging), residual genetic variation and newly acquired (epi)genetic variation could be experimentally distinguished in certain setups. If plants of single accessions are split up into populations that are maintained throughout the experiment, it is important to grow and phenotype them in a common environment both at the beginning and at the end of the experiment. If phenotypic differences between replicate populations of single accessions already occurred before imposing the selection treatments, it would suggest that the populations were (epi)genetically distinct from each other. However, if phenotypic differences between replicate populations emerge after the treatment and are maintained in a common environment, then this indicates that variation was gained during the treatment. We note that differences between replicate populations

due to founder effects or drift may be as relevant regarding evolutionary processes as are consistent differences between selection treatments (34).

We found weaker evidence for selection of epigenetic variation in this *Arabidopsis* experiment than we did in our previous study, in which we experimentally simulated selection in a rapidly changing environment (4). The reason for this difference may be that, in the previous study, we used RILs that were generated by crossing two different accessions, followed by several generations of inbreeding. We hypothesized that the epialleles differing between selected and ancestral plants persisted for multiple generations in a state that did not match the underlying genotype but finally converged to it (4). This scenario was impossible in the present study because we used distinct accessions. In our previous study (4), we also noted some residual heterozygosity in one of the RILs, which might have contributed to our observations. Studies on a wide range of plant species that included genetically non-uniform populations consistently reported a significant correlation between genetic and epigenetic variation (8, 9, 10, 11, and 12). To the best of our knowledge, there are very few studies in natural systems that demonstrate genetically independent, heritable — that is "true" epigenetic — variation in the absence of environmental triggers or after accounting for maternal effects. For example, maternal effects may explain the differences in flowering time between genetically identical apomictic dandelions collected in the field (35). One example for inheritance of true epigenetic variation has recently been shown in clonally reproducing duckweed *Lemna minor* (36). The authors report an epigenetic memory effect of heat treatments with changes in DNA methylation in the CHG context. However, this might be due to differences in mitotic (i.e., clonal) *vs.* meiotic inheritance. While it has long been known that certain epigenetic states are faithfully inherited during mitosis (37, 38), recent evidence shows that meiosis is accompanied by epigenetic reprogramming in plants (39, 40, 41, 42, 43, 44). Thus, it may be interesting to study epigenetic inheritance in species that generally reproduce clonally through mitosis but distinguish between vegetative reproduction through multicellular meristems and apomictic reproduction through an unreduced, unrecombined egg cell (45, 46, 1).

Methods

Experimental background and material from a multigenerational selection experiment

The selection experiment from which the plants used in this study were derived, was fully described by Züst and colleagues (25). In brief, the experiment began with a collection of 27 *Arabidopsis* accessions. Seeds obtained from NASC were expected to represent genetically uniform accessions and were propagated for one generation. For the selection treatment, experimental populations with 20 seeds from each accession were planted in cages. 17 days after sowing, different aphid species or a mixture were released into the cages (none as a control). After 60 days, all plants in a cage were harvested and a population of 800 seeds was sown representing the next generation. Each population was

maintained over five generations of selection, with six replicate populations per selection treatment. At the end of the selection experiment, the genetic composition of the populations was assessed (25).

Experimental design

For the study in a common environment, we focused on three well-surviving accessions from the selection experiment: Ma-0, Wei-0, and Wil-3. The accessions Ma-0 and Wei-0 were chosen because they were the most frequent after the aphid selection treatments. Wil-3 was chosen because it is a trichome-less accession that persisted across all treatments. We further focused on three selection treatments: the two single-aphid treatments with *Bravicoryne brassicae* (BB) and *Liphaphis erysimi* (LE), and the control without aphid treatment. Additionally, we included offspring from the plants that were used to set up the original selection experiment, here referred to as "ancestrals". All selected plants were grown together for one generation to minimize confounding maternal effects. For each treatment and accession, we planned to use five mother plants from each of the six populations per selection treatment. However, due to the rarity of some accessions in certain populations, we supplemented with mothers from other populations of the same selection treatment to fill up missing plants (360 maternal seed families were planned but ultimately only 261 were used). We accounted for this in our analysis by using population identity (n = 6 replicates x 3 selection treatments + ancestrals = 19) and maternal plant identity (n = 261 < 5 mothers x 3 accessions x 18 populations for control and selection-treatment plants + 30 mothers x 3 accessions for ancestrals = 360) as random terms. The ancestrals did not stem from multiple populations, but we planned to use an equal amount of mother plants per accession (n = 30). After the first generation of the present experiment (the sixth generation (G6) in reference to the selection experiment), between 5 (planned) and 150 (if there was only a single mother available) offspring of each mother were grown and phenotyped in the second (G7) generation (n = 1,800). Both generations were grown in a common environment in a fully randomized block design. Hence, if there were differences between selection treatments, they would have been maintained for two generations in the absence of selection by aphids.

G6: genotyping of offspring from the selection experiment

In the selection experiment, seeds were collected from all plants per cage (= population) at once. Thus, the accessions were mixed within populations. To identify the desired accessions, 2,000 seeds were sterilized and sown on Murashige and Skoog (MS) agar plates and stratified for three days at 4 °C. Plates were transferred to an incubator with a day/night cycle of 18h/6h at a constant temperature of 22 °C. Seven days after germination, we harvested leaves from seedlings resembling the Ma-0, Wei-0, or Wil-3 accession. Seedlings were genotyped as described previously (25). Genotyped seedlings were transplanted onto soil 14 days after germination and were grown in a controlled climate chamber at 18° C under a day/night cycle of 16h/8h. Seeds of plants from this G6 generation were collected from each plant separately to start the G7 generation with one seed per plant (n = 1,800; 1,759 surviving).

G7: genotyping and phenotyping

The plants of the G7 generation were arranged in a randomized layout on MS agar plates and later in trays (24 plants per tray). We changed from square plates to circular plates after the first three blocks. This caused a change in the thickness of the MS agar on the plates and subsequently to slower growth in blocks 4 and 5 than in blocks 1 to 3. We accounted for these differences during statistical analyses by using a corresponding contrast term within the term blocks. Plants were transplanted onto soil 14 days after germination and grown in a controlled climate chamber at 18° C under a day/night cycle of 16h/8h. Plants were grown in five randomized blocks of 360 plants with offspring of each mother present in each block. Plants were not switched between blocks when transferred from agar to soil. The remaining seedling DNA, which was not used for genotyping, was stored at –20 °C for later sequencing experiments.

Several traits were measured during the G7 generation (cf. 4). The days to bolting was measured as an indication for flowering time. The number of rosette leaves and the rosette diameter (longest distance) were both assessed at the day of bolting. 21 days after germination, the fifth true leaf was removed and stuck onto a foil for trichome number counts and leaf blade area measurements. To count trichomes, leaves were imaged on a Leica DMR bright-field microscope. Leaf area was measured by scanning the foil with an office scanner and extracting the leaf area with the ImageJ plugin LeafJ (47).

Statistical analyses

Statistical models

All analyses were carried out in R Version 3.6.1. Raw data are available (Supplementary data S4). The design comprised the following factorial terms that were not of interest per se: accession (ET, 3 levels), block (BL, 5 levels with a fixed-effect contrast EX comparing blocks 1–3 with blocks 4–5), tray (TR, 75 levels), population identity (LI, 18 levels + 1 ancestral per accession), mother plant identity (MO, < 120 levels per accession, 261 in total). The factorial term of interest was the selection treatment (Selection, 4 levels), which we split into three orthogonal one-degree of freedom contrasts to compare (i) ancestral plants with plants from the selection experiment (AN), (ii) the control treatment with the two selection treatments (CO), and (iii) the two selection treatments with each other (SE). We used both, the regular lm() function with specifying the appropriate statistical tests manually and lmer() that handles linear mixed models with random terms (48). With the regular lm(), we used the following two formulae for the explanatory terms:

- (EX+BL)+TR+(AN+CO+SE+LI)+ET+(AN+CO+SE+LI):ET+MO (m1)
- (EX+BL)+TR+(AN+CO+SE+LI)+ET+EX:(AN+CO+SE+LI)+EX:ET+LI:ET+MO (m2)

The term selection and its contrasts were tested against the term population identity (LI). Interactions of selection contrasts with accession (e.g., SE:ET) or with the block contrast (e.g., SE:EX) were tested against the interaction between population identity and

accession (LI:ET). The accession term itself (ET) was tested against the term mother plant identity within accessions (MO). Both models were also fitted with lmer() using the following formulae:

- (AN+CO+SE)+ET+(AN+CO+SE):ET+(1|BL)+(1|TR)+(1|MO)+(1|LI/ET) (m1)
- EX+(AN+CO+SE)+ET+EX:(AN+CO+SE)+ET:EX+(1|BL)+(1|TR)+(1|MO)+(1|LI/ET) (m2)

The two approaches yielded practically identical results, and we refer to the results obtained with the classical lm() function in the main text. Results from the lmer() tests are given in Supplementary data S1.

We ran the models for all phenotypic traits separately. We transformed all traits using the logarithm and the square root, and tested whether the transformations improved the normality assumption for the residuals using a Shapiro-Wilk test (49). In general, transformations did not improve the normality assumption except for trichome density, which was then transformed using the logarithm. We further assessed the effects of outliers by removing potential outliers with the ROUT method using the accession and the selection treatment for grouping (50). Accession Wil-3 had only one LE and three BB mothers, meaning that the two selection treatments were poorly replicated for this accession. As Wil-3 lacks trichomes, it was excluded from the trichome density analysis. We also ran all the models without Wil-3. To demonstrate that we tested exhaustively, we report all our results.

Power estimations

To estimate the statistical power to detect selection effects within genotypes (i.e., accessions), we employed the same models as described above and simulated data using the same data structure as in the real data. We simulated five different settings with effect sizes ranging from 1% to 20%: (i) differences between all four groups with step sizes of 1%, 5%, 10%, 15%, and 20% between each group (with the level following AN<CO<BB<LE, note that the largest difference here is 60%); (ii) only different in ancestors; (iii) only different in controls; (iv) different in both selection treatments; (v) different in only one selection treatment. For each setting, we generated 1,000 data sets and employed the models as described above. We then summarized the percentage of significant outcomes for the different settings (Supplementary data S2). In brief, we could detect effect sizes of 10% in more than 80% of the simulated data sets. An exception was trichome density for which the effect size had to be 20% for a robust detection. Note that an effect size of 20% is generally considered small (51).

DNA methylation

Library preparation

We selected 7, 12, and 13 mother plants from the ancestral, BB, and LE treatments (accession Wei-0). Libraries for whole-genome bisulfite sequencing were prepared as previously described (4), using the DNA extracted for genotyping of the plants in the G6 generation and sequenced on an Illumina HiSeq 2500 (250 bp paired-end). In total, we

obtained 293,022,365 read-pairs. On average, samples had about 8.6 mio reads. The lowest were 3.8, 6.4, and 6.5 mio, the highest were 11.7, 12.8, and 13.8 mio reads. Mapping efficiencies were generally similar (between 37.8% and 48.2%) but two samples had clearly lower mapping efficiencies (6.0% and 13.4%). These two samples were removed from the analysis (A18 and B26).

Data processing

Short reads were (quality) trimmed with fastp (version 0.20.0 with the options --trim_front1 6 --trim_tail1 6 --trim_front2 6 --trim_tail2 6; 52) and aligned to the reference genome with Bowtie2 (version 2.3.5.1, 53) in combination with Bismark (version 0.22.3 with the option --dovetail, 54). The reference genome for accession Wei-0 was obtained from 1001genomes.org (GMI-MPI release v3.1, pseudogenomes, pseudo6979.fasta.gz) and plastid sequences were added from the Col-0 reference genome (TAIR10). Alignments were deduplicated with Bismark and methylation tables were extracted with MethylDackel (version 0.5.0, github.com/dpryan79/MethylDackel). Files were merged and only cytosines with a coverage between 5 and 100 within at least two samples per group were kept (434,671; 479,201; and 2,669,064 in the CG, CHG, and CHH context, respectively; in total about 8.5% of all available cytosines). The three experimental groups (ancestrals and *B. brassicae* or *L. erysimi* aphid treatments) were compared to each other with the R-package DSS (version 2.32.0, 55), using the functions DMLfit.multiFactor(), DMLtest.multiFactor(), and callDMR(). The P-value threshold for the DMR calling was set to 0.01. Finally, all *P*-values were corrected for multiple testing and thereby converted to false discovery rates (FDRs, 56).

To test whether the extent of epigenetic variation changed during the experiment, we compared variation within original replicate populations with at least 4 samples with the function betadisper() from the R-package vegan (version 2.5.7, 57, 9). The function was also used to extract the distances to the centroids. From the 32 samples, 28 were within original replicate populations with at least 4 samples: 7 ancestrals, 12 from the *L. erysimi* treatment (3 original replicate populations), and 9 from the *B. brassicae* treatment (2 original replicate populations). Epigenetic distances between samples were calculated as the average difference in percent methylation across all cytosines (8) that passed the following filter: coverage between 5 and 100 in at least 5 individuals per treatment (n = 901,843).

To functionally characterize the genes associated with differentially methylated cytosines, we tested for enrichment of gene ontology (GO) terms using topGO 2.20 (58) in conjunction with the GO annotation available through biomaRt (59). Analysis was based on gene counts (genes with DMCs within 2kb compared to all genes with one or more tested cytosines that had an average of at least 10% methylation, i.e., potential gbM genes) using the "weight" algorithm with Fisher's exact test (both implemented in topGO). A term was identified as significant within a given parameter combination if the *P*-value was below 0.01.

Analysis of genetic variation in the subset 190 plants studied by Lin and colleagues (26)

To assess whether there was any genetic variation associated with the treatments in the experiment described by Lin and colleagues (2024), we used their RNA-seq data (PRJNA997595) on a subset of 190 plants that supposedly included 4 accessions. To obtain DNA sequence variant data from RNA-seq data, we followed, as far as possible (i.e., base recalibration required known sites that were not available), the GATK-guidelines (https://gatk.broadinstitute.org/hc/en-us/articles/360035531192-RNAseq-short-variant-discovery-SNPs-Indels?id=3891). In brief, short reads were (quality) trimmed with fastp (version 0.23.4; 52) and aligned to the reference genome with STAR (version 2.7.11a with the options --alignIntronMax 10000 --alignMatesGapMax 10000 --outSAMstrandField intronMotif and supplying the sample ID as read-group information; 60). Only unique alignments with a quality of at least 250 were kept. Read duplicates were marked with Picard tools (version 2.18.25; broadinstitute.github.io/picard) and spliced reads were split with GATK SplitNCigarReads. Bam files were merged into a single file to reduce running time while calling variants. Data were subset to 10,000 randomly chosen genes. Variants were finally called with GATK HaplotypeCaller (with the options --dont-use-soft-clipped-bases --standard-min-confidence-threshold-for-calling 20.0) and filtered with GATK VariantFiltration (with the options --window 35 --cluster 3 --filter-name "FS" --filter "FS > 30.0" --filter-name "QD" --filter "QD < 2.0"). The resulting VCF file was then filtered for variants with a coverage between 10 and 1,000 within at least four plants per accession and treatment. Variants with a minor allele frequency below 5% were removed. 13,418 variants passed this filter. We did not filter accessions separately when plotting them individually to maintain the distance values. Genetic pairwise distances were calculated as the fraction of alleles that differed between two individuals (61). We visualized the distances as previously described (61). When we observed that some plants within two of the accessions were clearly distinct from each other, we used the data on more than 1000 *Arabidopsis* accessions available from 1001genomes.org (https://1001genomes.org/data/GMI-MPI/releases/v3.1/1001genomes_snp-short-indel_only_ACGTN.vcf.gz). We intersected the two data sets (12,686 variants remained) and calculated all pairwise distances between the experimental plants and the publicly available accessions. Experimental plants were then assigned to the accession that was closest to it (Supplementary data S3).

Acknowledgements

We thank Tobias Züst (University of Zurich) for providing the seeds which served as starting material for this study, Christian Heichinger (Roche) for advice on genotyping and experimental setup, and Christina Westermann (University of Zurich) for editing the manuscript; Vimal Rawat, Gregor Rot, and Deepak Kumar Tanwar (University of Zurich) for discussions; Valeria Gagliardini, Christof Eichenberger, Daniela Guthörl, Arturo Bolaños, and Peter Kopf (University of Zurich) for general lab support; and Karl Huwiler (University of Zurich) for plant care. This work was supported through core

funding of the University of Zurich and a project of the University Research Priority Program "Evolution in Action".

Additional Information

Author contributions

UG and BS conceived and supervised the project and raised funding. KK performed the experiments. MWS, KK and BS analyzed the data. SW contributed to data analysis. MWS, KK, and UG wrote the manuscript with input from the other authors. All authors read and approved the final manuscript.

Author ORCIDs

Marc W. Schmid: https://orcid.org/0000-0001-9554-5318

Klara Kropivšek: https://orcid.org/0000-0003-1866-4094

Samuel E. Wuest: https://orcid.org/0000-0003-3982-0770

Bernhard Schmid: https://orcid.org/0000-0002-8430-3214

Ueli Grossniklaus: https://orcid.org/0000-0002-0522-8974

Additional Files

Supplementary files

Supplementary data 1: Collection of workbooks with results from the different analyses. The files are named as <phenotype>_<data set>.xlsx. For each phenotype, there are analyses with all three accessions ("all"), only the two accessions with trichomes ("NoWil3"), or with only one accession ("OnlyWei0", "OnlyMa0", and "OnlyWil3"). For each of them, there is one version with all data included and one in which the extreme values were removed (".outliersRemoved"). For example: bladeArea_all.outliersRemoved.xlsx refers to the analysis of the blade area using all three accessions and removing extreme values. Within each workbook, there are three sheets for untransformed data ("y"), log-transformed data ("logy"), and square-root-transformed data ("sqrty"). Within each sheet, the output of the two models tested either with the regular linear model (lm) or the random effects model (lmer) are given. In case of the regular lm output, alternative *F*- and *P*-values (F_altern and P_altern) refer to the tests against the correct error stratum (i.e., nested model). The columns F and P contain the *F*- and *P*-values from tests against the residuals. Regarding the model terms: "__Blocks" is the five level factor of experimental blocks that was split into two contrasts ("EX" and "BL"); "__Selection" is the four-level factor with the selection treatments that was split into three contrasts ("AN", "CO", and "SE"). "__Selection_X_accession/experiment" is accordingly the interaction of the four-level factor with the selection treatments and the accession ("ET") or the "EX" contrast. Abbreviations: Df: Degrees-of-freedom, SS: Sum

of squares, MS: Mean squares, F: *F*-value, P: *P*-value, F_altern: *F*-value of nested model, P_altern: *P*-value of nested model, perc_SS: %-Sum of squares explained by the term.

Supplementary data 2: A workbook containing tables from the power estimations. For each phenotype, power was estimated using all three accessions or only the two accessions with trichomes. In case of trichome density, only the latter is available. Each table contains the percentage of significant outcomes for a given simulation and factor/contrast of interest (based on 1,000 simulations). The simulations are named as <type>_<effect size>. There were five different types of simulations with effect sizes ranging from 1% to 20%: ("step") differences between all four groups with step sizes of 1%, 5%, 10%, 15%, and 20% between each group (with the level following AN<CO<BB<LE, note that the largest difference here is 60%); ("onlyAn") only different in ancestors; ("onlyCo") only different in controls; ("onlySe") different in both selection treatments; ("onlyLe") different in only one selection treatment.

Supplementary data 3: Table with the samples from Lin and colleagues (26). Given are the sample IDs, line ID, treatment, generation, accession label (used by Lin et al., 2025), and the accession assignments based on the genotypes inferred from the RNA-seq data and the publicly available data from 1001 *Arabidopsis* accessions (this study).

Supplementary data 4: Table with the phenotype data. Rosette diameter in mm, blade area in mm^2. Factors are labelled as described in Materials and Methods.

Data availability

Accession codes: Short reads were deposited at the NCBI Sequence Read Archive (SRA, www.ncbi.nlm.nih.gov/sra) and are available through accession number PRJNA1103971. Phenotype data were deposited on Zenodo (DOI: XXX).

Figure legends

Figure 1: Schematic of the original experimental design (top) and the current study (bottom).

Figure 2: Phenotype measures in the G7 generation averaged by the ID of the maternal lines from the G6 generation. *P*-values for contrasts with $P < 0.05$ are indicated (based on the analyses done per accession). *P*-values with asterisk were not significant when outliers were removed (see Figure 3).

Figure 3: Phenotypes measured in the G7 generation. Brown and green colors represent plants from blocks 1–3 and 4–5, respectively (slower growth in blocks 4–5). Dots with black strokes were identified as outliers. The number of plants available for the test differed slightly between the measurements. The number of plants given at the bottom of the figure corresponds to the minimal number of plants (i.e., for each phenotype there were at least N_{plants} plants). It varies by 2–3 plants per accession and treatment as some measures were missing. *P*-values for contrasts with $P < 0.05$ are indicated (based on the analyses done per accession). *P*-values with asterisk were not significant when outliers were removed.

Figure 4: A) Location of the differentially methylated cytosines along the genome. B) Epigenetic variation within different populations (selection replicates), i.e., distances to each centroid. Only populations with at least four individuals are shown. No overall and no pairwise *P*-value was smaller than 0.05.

Figure 5: A) Genetic distances between the 190 RNA-seq samples from Lin and colleagues (26). The two genetic clusters diverging from the rest of their accession group are highlighted. B) Within accession-label genetic distances, drawn to scale. Maximal distance given below the label, colored according to treatment and generation. The two treatments that have been assigned the wrong accessions are highlighted. C) Pairwise genetic distances within lines (plants of the same lineage, i.e., single-seed descendants), within accessions, and between accessions using either the original accession labels from Lin and colleagues (left; 26) or the accessions assigned based on genetic similarity to the data available at 1001genomes.org (right). The red line demarcates the distance equal to the average of within-line distances plus three standard deviations.

Tables

	Blade area		Time to bolting		Rosette diameter		Num. rosette leaves	
Model term	%-SS	*P*-value	%-SS	*P*-value	%-SS	*P*-value	%-SS	*P*-value
Blocks	46.97	**<0.001**	6.55	**<0.001**	43.23	**<0.001**	27.1	**<0.001**
> Blocks 1-3 *vs.* blocks 4-5 (EX)	41.85	**<0.001**	0.51	**<0.001**	39.62	**<0.001**	26.42	**<0.001**
> Remaining blocks	5.12	**<0.001**	6.04	**<0.001**	3.61	**<0.001**	0.68	**<0.001**
Selection treatments (ST)	0.1	0.887	0.24	0.925	0.76	**0.033**	0.5	0.782
> Ancestral *vs.* rest (AN)	0.02	0.742	0.06	0.735	0.11	0.222	0.24	0.484
> Control *vs.* treated (CO)	0.07	0.495	0.05	0.757	0.33	**0.043**	0.21	0.506
> *B. brassicae vs. L. erysimi* (SE)	0.01	0.852	0.13	0.624	0.32	**0.047**	0.05	0.746
Original replicate populatiton (LI)	2.38	**<0.001**	8.82	**<0.001**	1.18	**<0.001**	7.88	**<0.001**
Accession (ET)	8.16	**<0.001**	23.23	**<0.001**	2.2	**<0.001**	14.7	**<0.001**
ET:ST	0.42	**0.026**	0.34	0.41	0.62	**0.002**	0.16	0.655

> ET:AN	0.05	0.315	0.09	0.427	0.35	**0.001**	0.12	0.238
> ET:CO	0.06	0.288	0.22	0.153	0.01	0.84	0.01	0.843
> ET:SE	0.31	**0.005**	0.03	0.776	0.26	**0.005**	0.03	0.707
ET:LI	0.37	0.443	0.88	0.076	0.3	0.864	0.66	0.175
Mothers	6.07	**0.007**	8.11	0.216	7.38	**0.024**	6.81	0.342

Table 1: Percent sum of squares explained by model terms and *P*-values using all three accessions. Contrasts of a term are marked with indents. Complete output and further results for individual accessions are given in the supplementary files.

	Blade area		Time to bolting		Rosette diameter		Num. rosette leaves		Trichome density	
Model term	%-SS	*P*-value	%-SS	*P*-value	%-SS	*P*-value	%-SS	*P*-value	%-SS	*P*-value
Blocks	48.54	**<0.001**	5.91	**<0.001**	42.45	**<0.001**	29.24	**<0.001**	28.97	**<0.001**
> Blocks 1-3 *vs.* blocks 4-5 (EX)	43.45	**<0.001**	0.28	**0.026**	38.73	**<0.001**	27.94	**<0.001**	25.28	**<0.001**
> Remaining blocks	5.09	**<0.001**	5.63	**<0.001**	3.72	**<0.001**	1.3	**<0.001**	3.69	**<0.001**
Selection treatments (ST)	0.06	0.926	0.53	0.827	0.19	0.613	0.19	0.941	0.18	0.901
> Ancestral *vs.* rest (AN)	0	0.955	0.17	0.597	0	0.888	0.11	0.641	0	0.936
> Control *vs.* treated (CO)	0	0.908	0.21	0.561	0.1	0.338	0.07	0.715	0.02	0.801
> *B. brassicae vs. L. erysimi* (SE)	0.06	0.515	0.15	0.628	0.09	0.367	0.01	0.874	0.16	0.49
Original replicate populatiton (LI)	2.16	**<0.001**	9.5	**<0.001**	1.67	**0.002**	7.96	**<0.001**	4.99	**<0.001**
Accession (ET)	5.78	**<0.001**	18.5	**<0.001**	2.46	**<0.001**	11.78	**<0.001**	13.98	**<0.001**
ET:ST	0.12	0.424	0.42	0.349	0.38	**0.036**	0.04	0.799	0.07	0.572
> ET:AN	0.02	0.471	0.19	0.225	0.13	0.06	0.02	0.51	0.04	0.304

> ET:CO	0.03	0.394	0.18	0.241	0.01	0.588	0.02	0.522	0.02	0.422
> ET:SE	0.07	0.222	0.05	0.515	0.24	**0.017**	0	0.775	0.01	0.66
ET:LI	0.34	0.241	0.89	**0.048**	0.21	0.775	0.3	0.609	0.28	0.597
Mothers	7.69	**0.023**	11.27	0.319	9.16	0.2	10.2	0.106	8.25	0.46

Table 2: Percent sum of squares explained by model terms and *P*-values using only accessions Ma-0 and Wei-0. Contrasts of a term are marked with indents. Complete output and further results for individual accessions are given in the supplementary files.

GO.ID	Term	Found	Expected	*P*-value
GO:0070646	protein modification by small protein removal	7	1.53	0.00083
GO:0016571	histone methylation	10	3.07	0.00101
GO:0008284	positive regulation of cell proliferation	5	0.8	0.00119
GO:0042732	D-xylose metabolic process	3	0.3	0.00321
GO:0006487	protein N-linked glycosylation	5	1.03	0.0036
GO:0071702	organic substance transport	25	11.89	0.00423
GO:0015691	cadmium ion transport	2	0.11	0.00455
GO:0009834	plant-type secondary cell wall biogenesis	3	0.35	0.00482
GO:0006306	DNA methylation	7	2.09	0.00492
GO:0006342	chromatin silencing	8	2.86	0.00794
GO:0071705	nitrogen compound transport	23	11.74	0.00931
GO:0050665	hydrogen peroxide biosynthetic process	4	0.83	0.00945
GO:1905039	carboxylic acid transmembrane transport	2	0.15	0.00945

Table 3: Gene Ontology (GO) enrichment among genes mapped by DMCs. "Found" indicates the number of different genes with the term that were mapped by DMCs. "Expected" refers to the number of genes that were expected to be mapped by DMCs if DMCs were randomly distributed.

References

1. Schmidt, A., Schmid, M. W. & Grossniklaus, U. Plant germline formation: Common concepts and developmental flexibility in sexual and asexual reproduction. *Development* **142,** 229–241 (2015).

2. Soppe, W. *et al.* The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Molecular Cell* **6,** 791–802 (2000).

3. He, L. *et al.* A naturally occurring epiallele associates with leaf senescence and local climate adaptation in *Arabidopsis* accessions. *Nature Commununications* **9,** 460 (2018).

4. Schmid, M. W. *et al.* Contribution of epigenetic variation to adaptation in *Arabidopsis*. *Nature Communications* **9,** 4446 (2018).

5. van der Graaf, A. *et al.* Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *Proceedings of the National Academy of Sciences U.S.A.* **112,** 6676–6681 (2015).

6. Johannes, F. & Schmitz, R. J. Spontaneous epimutations in plants. *The New Phytologist* **221,** 1253–1259 (2019).

7. Hazarika, R. R. *et al.* Molecular properties of epimutation hotspots. *Nature Plants* **8,** 146–156 (2022).

8. van Moorsel, S. J. *et al.* Evidence for rapid evolution in a grassland biodiversity experiment. *Molecular Ecology* **28,** 4097–4117 (2019).

9. Mounger, J. *et al.* Inheritance of DNA methylation differences in the mangrove *Rhizophora mangle*. *Evolution & Development* **23,** 351–374 (2021).

10. Mounger, J. M. *et al.* Genetic and epigenetic differentiation across intertidal gradients in the foundation plant *Spartina alterniflora*. *Frontiers in Ecology and Evolution* **10,** 868826 (2022).

11. Boquete, M. T. *et al.* Molecular basis of intraspecific differentiation for heavy metal tolerance in the copper moss *Scopelophila cataractae*. *Environmental and Experimental Botany* **201,** 104970 (2022).

12. Ibañez, V. N. *et al.* Environmental and genealogical effects on DNA methylation in a widespread apomictic dandelion lineage. *Journal of Evolutionary Biology* **36,** 663–674 (2023).

13. Dubin, M. J. *et al.* DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaption. *eLIFE* **4,** e05255 (2015).

14. Kawakatsu, T. *et al.* Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions. *Cell* **166,** 492–505 (2016).

15. Cortijo, S. *et al.* Mapping the epigenetic basis of complex traits. *Science* **343,** 1145–1148 (2014).

16. Kooke, R. *et al.* Epigenetic basis of morphological variation and phenotypic plasticity in *Arabidopsis thaliana*. *The Plant Cell* **27,** 337–348 (2015).

17. Paszkowski, J. & Grossniklaus, U. Selected aspects of transgenerational epigenetic inheritance and resetting in plants. *Current Opinion in Plant Biology* **14,** 195–203 (2011).

18. Jablonka, E. & Raz, G. Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly Review of Biology* **84,** 131–176 (2009).

19. Quadrana, L. & Colot, V. Plant transgenerational epigenetics. *Annual Review of Genetics* **50,** 467–491 (2016).

20. Richards, C. L. *et al.* Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecology Letters* **20,** 1576–1590 (2017).

21. Richards, E. Natural epigenetic variation in plant species: A view from the field. *Current Opinion in Plant Biology* **14,** 204–209 (2011).

22. Preite, V. *et al.* Increased transgenerational epigenetic variation, but not predictable epigenetic variants, after environmental exposure in two apomictic dandelion lineages. *Ecology and Evolution* **8,** 3047–3059 (2018).

23. Sammarco, I. *et al.* DNA methylation in the wild: Epigenetic transgenerational inheritance can mediate adaptation in clones of wild strawberry (*Fragaria vesca*). *The New Phytologist* **241,** 1621–1635 (2024).

24. Greaves, I. *et al.* Epigenetic changes in hybrids. *Plant Physiology* **168,** 1197–1205 (2015).

25. Züst, T. *et al.* Natural enemies drive geographic variation in plant defenses. *Science* **338,** 116–119 (2012).

26. Lin, X. *et al.* Environment-induced heritable variations are common in *Arabidopsis thaliana*. *Nature Communications* **15,** 4615 (2024).

27. Kartal, Ö., Schmid, M. W. & Grossniklaus, U. Cell type-specific genome scans of DNA methylation divergence indicate an important role for transposable elements. *Genome Biology* **21,** 172 (2020).

28. Bewick, A. J. & Schmitz, R. J. Gene body DNA methylation in plants. *Current Opinion in Plant Biology* **36,** 103–110 (2017).

29. Muyle, A. M., Seymour, D. J., Lv, Y., Huettel, B. & Gaut, B. S. Gene body methylation in plants: Mechanisms, functions, and important implications for understanding evolutionary processes. *Genome Biology and Evolution* **14,** evac038 (2022).

30. *Epigenetic mechanisms of gene regulation*. (Cold Spring Harbor Laboratory Press, 1996).

31. Becker, C. *et al.* Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480,** 245–249 (2011).

32. Schmitz, R. J. *et al.* Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334,** 369–373 (2011).

33. Atwell, S. *et al.* Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465,** 627–631 (2010).

34. *An introduction to population genetics theory*. (Harper; Row, 1970).

35. Wilschut, R., Oplaat, C., Snoek, B., Kirschner, J. & Verhoeven, K. Natural epigenetic variation contributes to heritable flowering divergence in a widespread asexual dandelion lineage. *Molecular Ecology* **25,** 1759–1768 (2016).

36. Van Antro, M. *et al.* DNA methylation in clonal duckweed (*Lemna minor* L.) lineages reflects current and historical environmental exposures. *Molecular Ecology* **32,** 428–443 (2023).

37. Meins, F. J. Heritable variation in plant cell culture. *Annual Review of Plant Physiology* **34,** 327–346 (1983).

38. Meins, F. J. Determination and morphogenetic competence in plant tissue culture. *Botanical Monographs* **23,** 7–25 (1986).

39. Slotkin, R. *et al.* Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell* **136,** 461–472 (2009).

40. She, W. & Baroux, C. Chromatin dynamics during plant sexual reproduction. *Frontiers in Plant Science* **5,** 354 (2014).

41. She, W. & Baroux, C. Chromatin dynamics in pollen mother cells underpin a common scenario at the somatic-to-reproductive fate transition of both the male and female lineages in *Arabidopsis*. *Frontiers Plant Science* **6,** 294 (2015).

42. Park, K. *et al.* DNA demethylation is initiated in the central cells of *Arabidopsis* and rice. *Proceedings of the National Academy of Sciences U.S.A.* **113,** 15138–15143 (2016).

43. Schmid, A., M.W. Giraldo-Fonseca, Rövekamp, M., Smetanin, D. & Grossniklaus, U. Extensive epigenetic reprogramming during the life cycle of *Marchantia polymorpha*. *Genome Biology* **19,** 9 (2018).

44. Borg, M. *et al.* Epigenetic reprogramming rewires transcription during the alternation of generations in *Arabidopsis*. *eLife* **10,** e61894 (2021).

45. Schmid, B. Some ecological and evolutionary consequences of modular organization and clonal growth in plants. *Evolutionary Trends in Plants* **4,** 25–34 (1990).

46. Koltunow, A. & Grossniklaus, U. Apomixis: A developmental perspective. *Annual Review of Plant Biology* **54,** 547–574 (2003).

47. Maloof, J. N., Nozue, K., Mumbach, M. R. & Palmer, C. M. LeafJ: An ImageJ plugin for semi-automated leaf shape measurement. *Journal of Visualized Experiments* **71,** e50028 (2013).

48. Schmid, B., Baruffol, M., Wang, Z. & Niklaus, P. A. A guide to analyzing biodiversity experiments. *Journal of Plant Ecology* **10,** 91–110 (2017).

49. Shapiro, S. S. & Wilk, M. B. An analysis of variance test for normality (complete samples). *Biometrika* **52,** 591–611 (1965).

50. Motulsky, H. J. & Brown, R. E. Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics* **7,** 123 (2006).

51. Cohen, C. *Statistical power analysis for the behavioral sciences*. (Lawrence Erlbaum Associates, Publishers, 1988).

52. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34,** i884–i890 (2018).

53. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with bowtie 2. *Nature Methods* **9,** 357–359 (2012).

54. Krueger, F. & Andrews, S. Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27,** 1571–1572 (2011).

55. Park, Y. & Wu, H. Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics* **32,** 1446–1453 (2016).

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

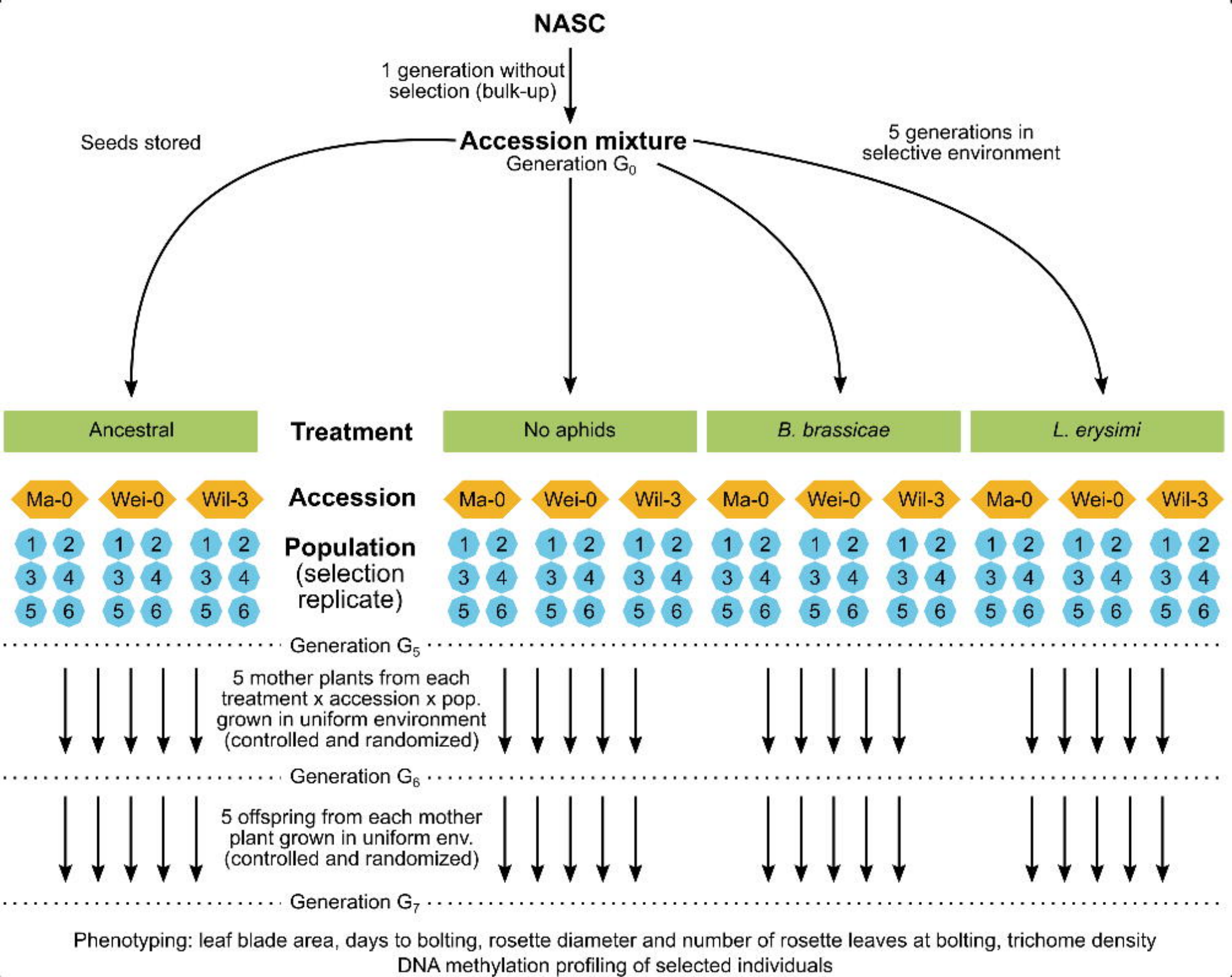
57. Dixon, P. VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science* **14,** 927–930 (2003).

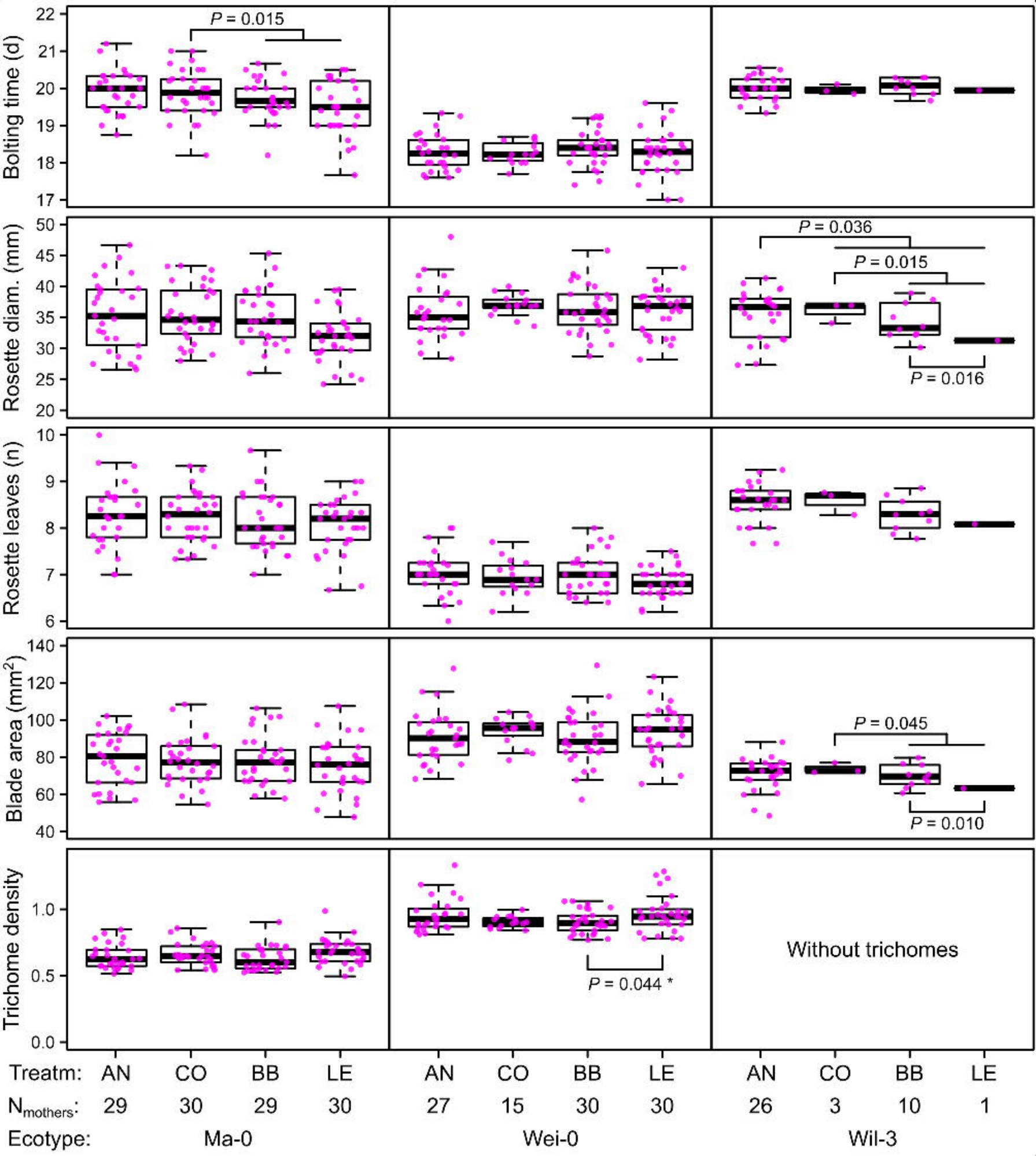
58. Alexa, A., Rahnenführer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22,** 1600–1607 (2006).

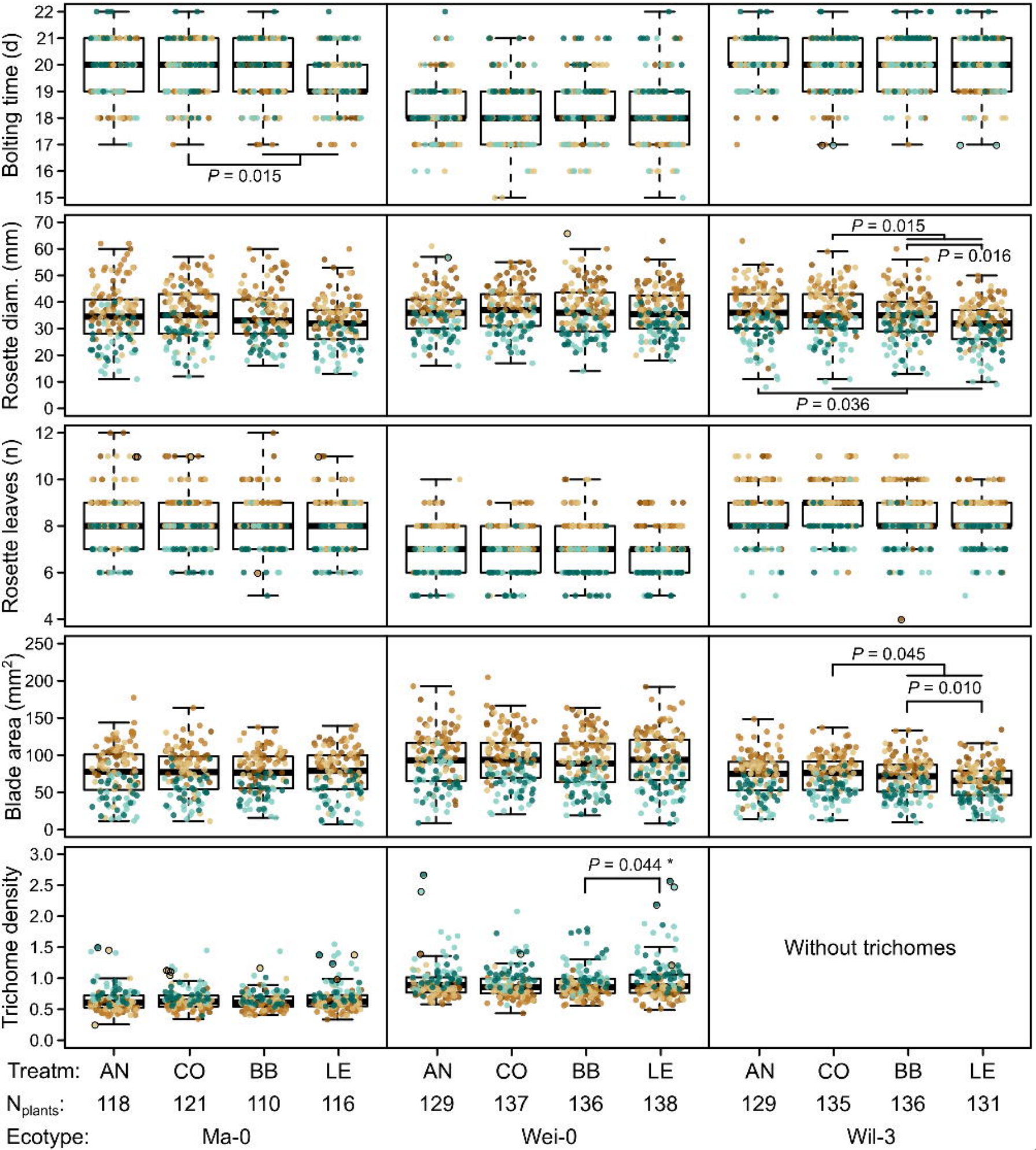
59. Durinck, S., Spellman, P., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols* **4,** 1184–1191 (2009).

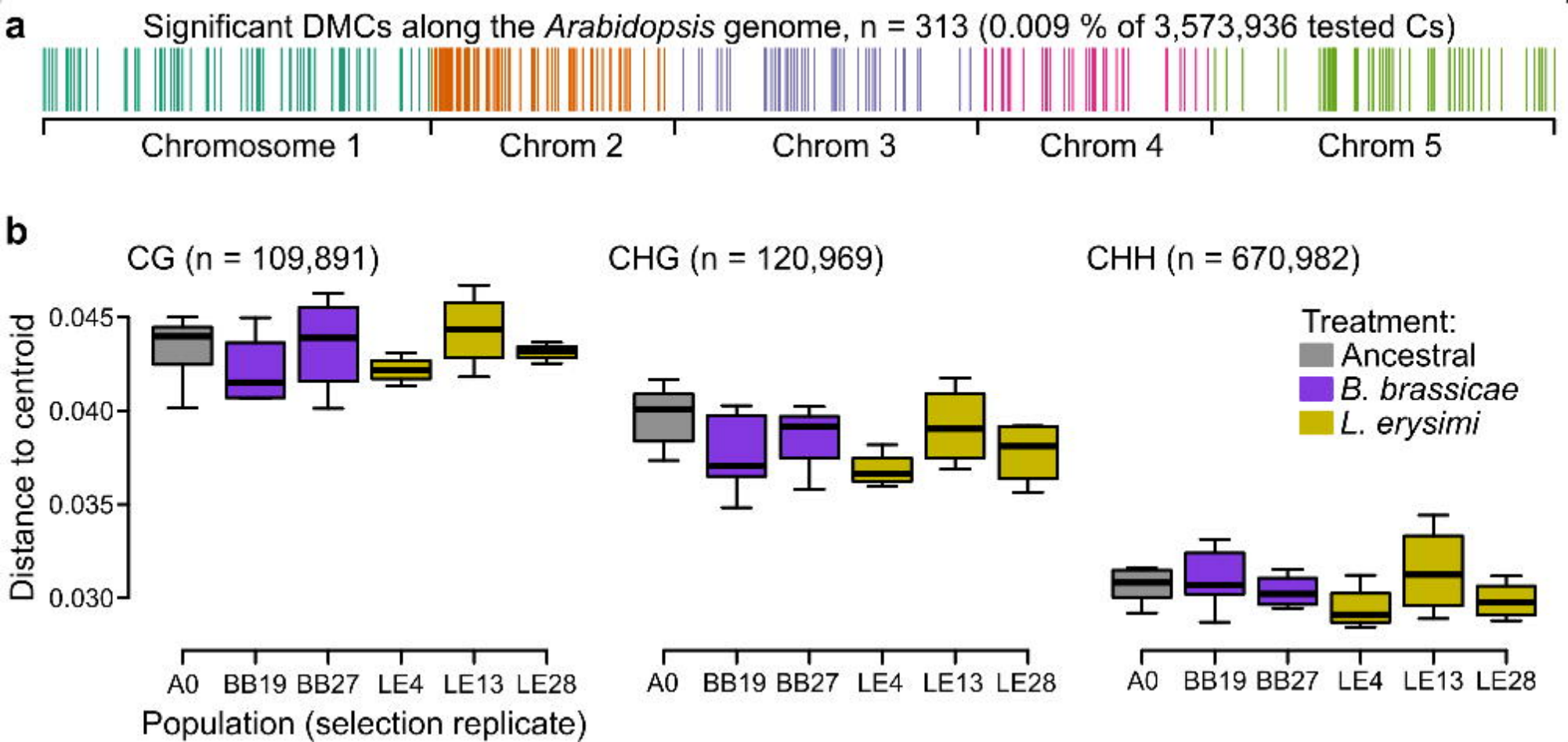
60. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).

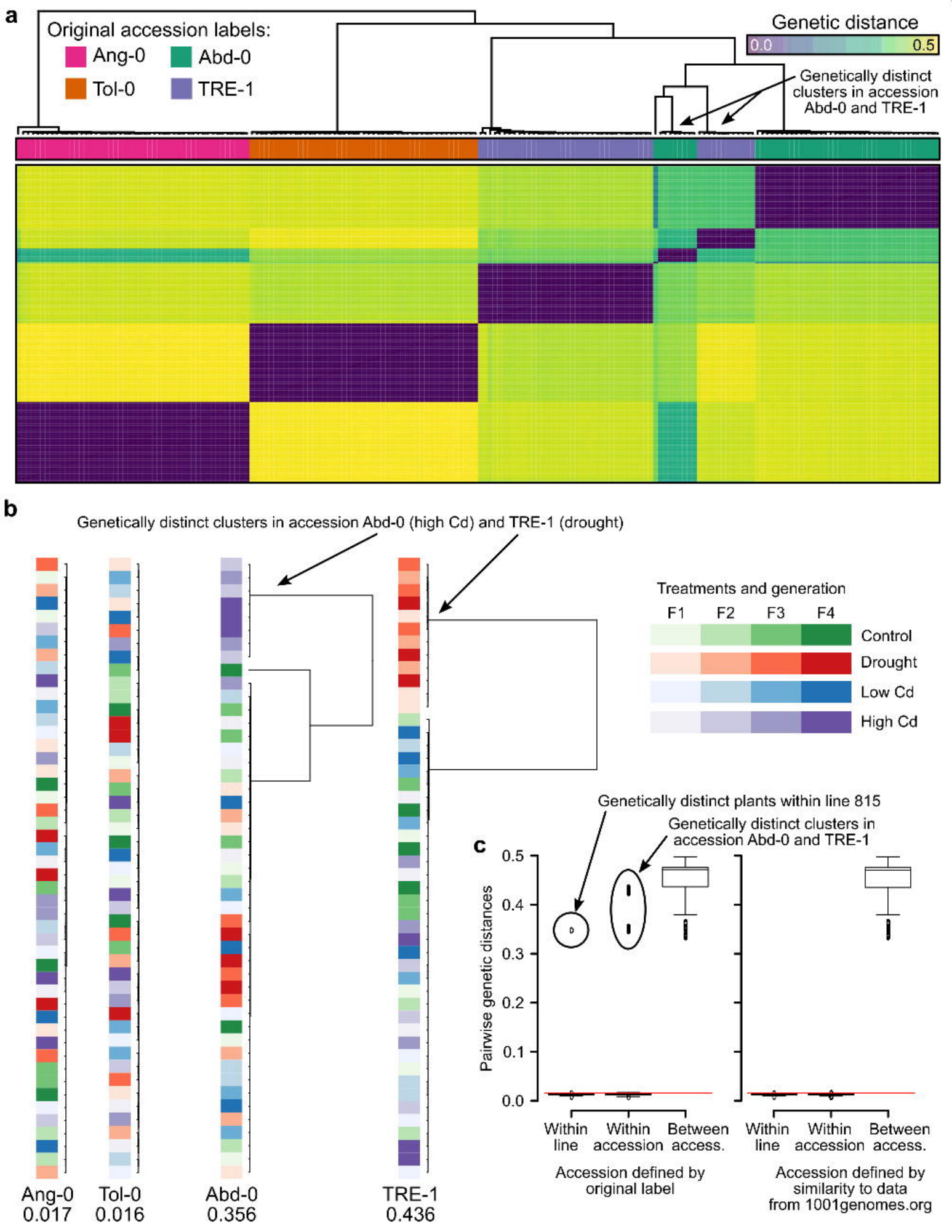
61. Schmid, M., Moradi, A., Leigh, D., Schuman, M., van Moorsel, S. Covering the bases: Population genomic structure of *Lemna minor* and the cryptic species *L. japonica* in Switzerland. *Ecology and Evolution* 14, e11599 (2024).


NASC
1 generation without selection (bulk-up)
Seeds stored
Accession mixture
Generation G_0
5 generations in selective environment
Ancestral
Treatment
No aphids
B. brassicae
L. erysimi
Ma-0
Wei-0
Wil-3
Accession
Population (selection replicate)
1 2 3 4 5 6
Generation G_5
5 mother plants from each treatment x accession x pop. grown in uniform environment (controlled and randomized)
Generation G_6
5 offspring from each mother plant grown in uniform env. (controlled and randomized)
Generation G_7
Phenotyping: leaf blade area, days to bolting, rosette diameter and number of rosette leaves at bolting, trichome density
DNA methylation profiling of selected individuals


Bolting time (d)
P = 0.015
Rosette diam. (mm)
P = 0.036
P = 0.015
P = 0.016
Rosette leaves (n)
Blade area (mm²)
P = 0.045
P = 0.010
Trichome density
P = 0.044 *
Without trichomes
Treatm: AN CO BB LE AN CO BB LE AN CO BB LE
N_mothers: 29 30 29 30 27 15 30 30 26 3 10 1
Ecotype: Ma-0 Wei-0 Wil-3


Bolting time (d)
P = 0.015
Rosette diam. (mm)
P = 0.015
P = 0.016
P = 0.036
Rosette leaves (n)
Blade area (mm²)
P = 0.045
P = 0.010
Trichome density
P = 0.044 *
Without trichomes
Treatm: AN CO BB LE AN CO BB LE AN CO BB LE
N_plants: 118 121 110 116 129 137 136 138 129 135 136 131
Ecotype: Ma-0 Wei-0 Wil-3


a
Significant DMCs along the Arabidopsis genome, n = 313 (0.009 % of 3,573,936 tested Cs)
Chromosome 1
Chrom 2
Chrom 3
Chrom 4
Chrom 5
b
CG (n = 109,891)
CHG (n = 120,969)
CHH (n = 670,982)
Distance to centroid
0.045
0.040
0.035
0.030
Treatment:
Ancestral
B. brassicae
L. erysimi
A0
BB19
BB27
LE4
LE13
LE28
Population (selection replicate)


a
Original accession labels:
Ang-0
Abd-0
Tol-0
TRE-1
Genetic distance
0.0
0.5
Genetically distinct clusters in accession Abd-0 and TRE-1
b
Genetically distinct clusters in accession Abd-0 (high Cd) and TRE-1 (drought)
Treatments and generation
F1
F2
F3
F4
Control
Drought
Low Cd
High Cd
Ang-0
0.017
Tol-0
0.016
Abd-0
0.356
TRE-1
0.436
c
Genetically distinct plants within line 815
Genetically distinct clusters in accession Abd-0 and TRE-1
Pairwise genetic distances
0.5
0.4
0.3
0.2
0.1
0.0
Within line
Within accession
Between access.
Within line
Within accession
Between access.
Accession defined by original label
Accession defined by similarity to data from 1001genomes.org