Local trade-offs shape flower size evolution across Arabidopsis thaliana distribution

Kévin Sartori, Clàudia F Mestre, Md Jonaid Hossain, Aurélien Estarague, Elza Gaignon, Pierre Moulin, Jesse R Lasky, Denis Vile, François Vasseur, Cyrille Violle, et al.

Local trade-offs shape flower size evolution across *Arabidopsis thaliana* distribution

Kevin Sartori[1*], Clàudia F. Mestre[1], Md Jonaid Hossain[1], Aurélien Estarague[2], Elza Gaignon[2], Pierre Moulin[2], Jesse R. Lasky[3], Denis Vile[4], Francois Vasseur[2,4], Cyrille Violle[2] and Adrien Sicard[1*].

[1] Department of Plant Biology, Swedish University of Agricultural, Sciences and Linnean Center for Plant Biology, Uppsala, Sweden.

[2] CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

[3] Department of Biology and PAC Herbarium, Pennsylvania State University, University Park, PA, USA

[4] Laboratoire d'Ecophysiologie des Plantes sous Stress Environnementaux (LEPSE), INRA, Montpellier SupAgro, UMR, Montpellier, France.

* Corresponding authors: kevin.sartori@slu.se ; adrien.sicard@slu.se

Author Contributions: `

Conceptualization: KS, FV, CV and AS; investigations: KS, CFM, MJH, EG, PM and AE; Data analyses: KS and CFM; Materials and analysis tools contributions: DV and JRL; writing original draft KS and AS; writing –review & editing– All authors.

Competing Interest Statement:

The authors declare no competing interests.

Keywords: Flower size evolution, *Arabidopsis thaliana*, phenotypic diversity, environmental constraints.

Abstract

The diversification of flowers, largely driven by mutualistic interactions with animal pollinators, has generated remarkable variation in floral form and size and is thought to have driven the evolutionary radiation of angiosperms. Here, we investigate the geographic variation in reproductive organ growth in the self-fertilizing species *Arabidopsis thaliana*, where the loss of pollinator dependence is predicted to favour reduced floral investment through resource reallocation. We find extensive variation in flower size underpinned by a polygenic architecture, in which derived alleles both increase and decrease petal size and show signatures of positive selection. The direction of flower-size evolution varies geographically, reflecting environmentally mediated shifts in the relationship between flower size and seed production. Strong purifying selection at the species' climatic margins favors smaller flowers, whereas relaxed environmental constraints in more suitable habitats allow the persistence and potential adaptive value of large-flower variants. This biogeographic pattern extends to other life-history traits, illustrating how environmental heterogeneity, acting through resource allocation constraints, shapes evolutionary trajectories and maintains phenotypic diversity in selfing lineages.

Introduction

Natural selection balances flower organ construction and maintenance costs against the likelihood of fertilization (Roddy et al., 2020). These costs follow classical allocation theory, as illustrated by the trade-off between flower size and number (Roddy et al., 2020)(Caruso, 2006; Worley & Barrett, 2000). In animal-pollinated outcrossing species, pollinator-mediated selection often favors large and more conspicuous flowers adapted to pollinator nutritional needs and behaviours, promoting flower diversification (Crane et al., 1994; Friis et al., 1994, 1999, 2006; Galen, 1999). Following the transitions to self-fertilization, however, selection is expected to reduce investment in floral displays and reallocate resources to other fitness components (Stebbins, 1957). This mechanism is frequently invoked to explain the characteristic reduction in flower size and attractiveness, collectively known as the "selfing syndrome", in selfing species (Sicard et al., 2011). Yet, empirical studies quantifying the relative importance of resource-allocation trade-offs compared to other evolutionary processes remain limited.

Different evolutionary processes could influence floral evolution in selfing species, each leading to distinct patterns of variation in flower size. When the construction of floral organs incurs significant energetic costs, selection should favor mutations that reduce floral display size, thereby enhancing seed production through resource reallocation (Charlesworth & Charlesworth, 1981). This directional selection scenario predicts (i) a negative association between flower size and seed production, (ii) genomic evidence of positive selection on alleles that reduce flower size, and (iii) a general convergence toward smaller flowers with reduced within-species variance (Stephan, 2019). Because environmental heterogeneity influences both resource availability and selection strength (Troth et al., 2018), these patterns should be more pronounced in resource-limited environments (Fu et al., 2010; MacTavish & Anderson, 2020; Tuller et al., 2018). Alternatively, the loss of attractive traits could confers little to no fitness benefit and selection maintaining large displays would be relaxed. Under this scenario the trajectory of floral-size evolution may become less predictable. In this case, the role of genetic drift should play a more dominant role, leading to idiosyncratic shifts in mean size, widening phenotypic variance, and few clear signatures of directional selection at causal loci (Woźniak & Sicard, 2018). Reduced effective recombination under high homozygosity can further facilitate drift-mediated fixation of nonsynonymous mutations that impair floral-growth genes (Barrett et al., 2014).

Additional processes may interact with these trajectories. Pleiotropy can couple floral size variation with other organs controlled by shared developmental pathways, producing correlated responses that range from near-isometric to clearly allometric depending on the extent of shared genetic control (Huxley et al., 1993; Johnson & Lenhard, 2011) . Likewise, linked selection, enhanced in selfers due to low effective recombination, can further reshape diversity around selected loci and indirectly hitchhike alleles affecting floral size (Slotte, 2014). Given the diversity and nature of the mechanisms involved, floral-size variation is likely to be highly polygenic, consistent with observations in selfing lineages (Abraham et al., 2013; Sicard et al., 2011; Sicard & Lenhard, 2011; Woźniak et al., 2020). However, the relative contributions of these mechanisms remain poorly understood.

Arabidopsis thaliana provides a powerful system to test these predictions. The evolution of self-fertilization in this species occurred relatively recently and was followed by rapid range expansion across diverse climates and environments (Bechsgaard et al., 2006; Shimizu & Tsuchimatsu, 2015; Tang et al., 2007). This history provides opportunities for environment-specific selection on floral investment and the fixation of distinct size-affecting mutations across lineages (Tsuchimatsu et al., 2020). Leveraging extensive genomic and phenotypic resources, we examine the genetic and ecological drivers of floral-size evolution following the emergence of selfing and assess how environmental gradients shape these evolutionary trajectories across the species' range. Our results reveal that patterns of phenotypic evolution differ geographically, reflecting variation in selection pressures across environments.

Results

Flower size is highly variable within A. thaliana

The directional selection hypothesis predicts that both the mean and variance of flower size decline in self-fertilizing species such as *Arabidopsis thaliana* due to resource allocation trade-offs that favour reduced investment in floral display. Alternatively, genetic and phenotypic variation may persist if drift predominates (Buysse et al., 2024) or if spatially variable environments maintain divergent selection (MacTavish & Anderson, 2020).

To test these predictions, we measured the length, width, and area of floral organs and leaves in 407 *A. thaliana* accessions spanning the species' global genetic and geographic diversity (**Fig. S1**; **Dataset S1**). All traits were variable and approximately normally distributed (**Fig. S2**). Floral organ size explained most phenotypic variance among accessions (43.2%, PC1; **Fig. 1a**), with strong correlations among floral traits ($r > 0.5$). Leaf size variation dominated PC2 (21.6% of the variance). Ovule number contributed mainly to PC4 (<8% of variance). Petal area showed the greatest variability, matching leaf area in its coefficient of variation and spanning more than a threefold range (**Dataset S2**).

These results indicate that floral-size variation in *A. thaliana* is substantial despite selfing, consistent with either incomplete fixation of size-reducing alleles during the evolution of selfing or environmentally structured selection maintaining divergence among lineages.

Independent, low-pleiotropy mutations largely account for the variation in flower organ size.

We investigated whether indirect effects, such as linked selection (Smith & Haigh, 1974) or pleiotropy with other traits under selection, drive the evolution of floral size in *Arabidopsis thaliana*. To this end, we first tested for phenotypic covariation between floral organ size and other traits among genotypes (**Dataset S2**). Stamen, petal, and sepal size were significantly and positively correlated ($r = 0.28–0.93$) but showed no correlation with ovule number or leaf size, except for a weak negative correlation between leaf length and petal width ($r = –0.15$; Fig. 1b). The relationship between long-stamen length and petal length was strictly isometric (slope = 1, P-value > 0.01), suggesting an overlap in their genetic

bases, while other pairwise relationships were allometric (slope < 1, P-value < 0.001). All floral organs were positively correlated with flowering time (r = 0.14–0.42). We further examined correlations with published phenotypes. Floral organ size traits showed no significant correlation with any of the 107 phenotypes from Atwell et al. (2010) (**Dataset S2**). Specific leaf area (SLA), leaf nitrogen content (LNC), and total plant biomass (Przybylska et al. 2023) were significantly correlated with several floral size traits (r = –0.35–0.4), and most traits were also correlated with flowering time (r = –0.67–0.70) in this dataset. To determine whether phenotypic correlations result from shared genetic bases, we quantified genetic correlations using multivariate GWAS (mGWAS) for significantly correlated trait pairs. The mGWAS partitions genetic variance into trait-specific and shared components, allowing estimation of the genetic correlation (r_g) between traits (Zhou & Stephens, 2014) (**Dataset S3**). Within our dataset, we detected no significant genetic correlations between floral and leaf traits, indicating independent genetic bases for their size evolution. Petal and sepal area showed weak but significant correlations with LNC and plant dry mass (r_g ~ 0.3, P < 0.01), proxies for resource acquisition ability (Pérez-Harguindeguy et al., 2013). These results suggest that resource acquisition capacity modestly influences both vegetative carbon economy and floral organ growth, supporting a role for allocation constraints in flower-size evolution. Overall, floral organ size variation showed limited phenotypic and genetic correlation with other traits, indicating that pleiotropy and linked selection exert only minor influence on floral size evolution in *A. thaliana*.

We next conducted univariate GWAS to identify loci underlying individual trait variation (**Fig. 1c**). Despite high SNP-based heritability (PVE = 0.55–0.85) across all floral traits, only three genome-wide significant peaks surpassed the Bonferroni threshold for sepal, petal, and stamen size, consistent with a largely polygenic basis. Two peaks were organ-specific, while one was shared between petal and sepal traits (**Fig. 1c**). To capture additional polygenic signals, we applied a relaxed significance threshold ($P < 5.2*10^{-5}$; see Methods) based on established associations. After filtering for linkage, we identified 61, 105, 87, 59, and 38 SNPs, corresponding to 144, 191, 194, 126, and 117 candidate genes, associated with leaf, petal, sepal, stamen size, and ovule number, respectively (**Fig. 1d**; **Dataset S4**). Functional annotation using TAIR gene ontology terms (Huala et al., 2001) identified several known growth regulators among candidate genes, including *JAGGED*, *GROWTH-REGULATING FACTOR 6*, *CYTOCHROME P450 90B1*, *KINASE-INDUCIBLE DOMAIN INTERACTING 9*, *E2FE*, *MYB DOMAIN PROTEIN 104*, *RESPONSE REGULATOR 7*, *INDOLE-3-ACETIC ACID INDUCIBLE 19*, and *SEEDSTICK* (Sauret-Güeto et al., 2013; Omidbakhshfard et al., 2015; Gómez-Ocampo et al., 2023; Gonzalez et al., 2015; Vlieghe et al., 2005; Dubos et al., 2010; Lee et al., 2008; Maki et al., 2022; Lopes et al., 2023). Genes associated with petal area were significantly enriched for functions related to floral and shoot development (6 genes, P ~0.01) and reproductive regulation (12 genes, P ~ 0.02; Dataset S5). We observed little overlap between SNPs associated with floral and leaf size, consistent with the absence of phenotypic and genetic correlations (**Fig. 1d**). A few genes were associated with both leaf and sepal growth through distinct SNPs, suggesting organ-specific effects of these mutations on general growth regulator activities. Despite moderate phenotypic correlations among floral organs, their genetic determinants showed limited overlap, consistent with low pleiotropy. Thus, floral organ size

appears to evolve largely independently, with observed phenotypic covariation more likely reflecting coordinated responses to shared environmental conditions rather than shared genetic architecture.

Flower-size genes are under a mixed selection pattern in Arabidopsis thaliana.

The reduction in floral organ size accompanying the transition to self-fertilization is thought to result from a reallocation of resources no longer required for pollinator attraction toward other fitness-enhancing traits. Under this hypothesis, new mutations reducing flower size should be selectively favored and accumulate in selfing lineages. To test this, we estimated the total effect size and ancestral origin of alleles for all associated SNPs (see Methods). More than three-quarters of derived alleles increased petal area (**Fig. 2a**). This excess of positive effects among derived alleles was unique to petal area (**Fig. 2b**).

To assess whether the fixation of derived alleles increasing flower size could result from genetic drift, we first computed genome-wide πN/πS, which contrasts the accumulation of non-synonymous (πN) and synonymous (πS) diversity to infer the strength and direction of selection. Genes containing SNPs associated with petal area exhibited a skewed πN/πS distribution that tended to fall below the genome-wide average (P-value < 0.1), indicating ongoing purifying selection rather than relaxed selection (**Fig. 2c**). Other floral organ size traits showed no comparable pattern (**Fig. S3**). While most petal candidate genes had πN/πS values below the genomic average, three genes, *AT1G36940*, *KIX9* and *FLM*, displayed unusually high values within the top 5% of the genome-wide distribution, suggesting diversifying selection on their protein sequences.

We next examined whether derived alleles in petal area genes show evidence of positive selection. At a bi-allelic site, alleles under selection are expected to be associated with longer haplotypes due to incomplete selective sweeps. We estimated haplotype decay around each focal SNP using genome-wide extended haplotype homozygosity (EHH) (Klassmann & Gautier, 2022). The standardized ratio of derived to ancestral haplotype homozygosity (the integrated haplotype homozygosity score, iHS) identifies alleles with unusually long haplotypes for their frequency. Two petal area SNPs exhibited high |iHS| values, resulting from an extended haplotype homozygosity around the derived alleles, a sign ofrecent positive selection. The first, snp_1_10937139 (P = 0.013), increased petal area and likely affects the growth regulator JMJ18. The second, snp_5_870273 (P = 0.02), decreased petal area and likely influences AT5G03480, a poorly characterized RNA-binding protein expressed during early petal development. A third SNP, snp_1_28960616 (P = 0.07), showed a marginal signal of positive selection and likely affects FLM, a known flowering-time regulator that also influences flower development (**Fig. 2d**) (Posé et al., 2013). As a control, we repeated the analysis for all floral and leaf traits and quantified the enrichment of trait-associated SNPs within the top 0.05% of the genome-wide iHS distribution (Tsuchimatsu et al., 2020). Leaf size–associated SNPs showed a threefold enrichment, whereas no enrichment was detected for flower size (**Fig. S4**).

Together, these results indicate a lack of consistent directional selection acting on the genetic determinants of flower size in *A. thaliana*. Neither purifying nor positive selection acts uniformly across petal size regulators, and no consistent directional trend is apparent. The observed variation likely

reflects spatially heterogeneous selection combined with neutral processes, accentuated by the species' small effective population size and strong population structure. We next investigated whether phenotypic and genetic variation in floral organ size arises from spatially heterogeneous selective pressures across *A. thaliana*'s range.

The influence of flower size on individual's performance is environment dependent

The finding that alleles increasing flower size can be under positive selection in a selfing species challenges the resource allocation theory, which predicts higher fitness associated with reduced investment in floral organ growth in species such as *A. thaliana*. To explore this apparent contradiction, we examined the relationship between petal size variation and plant fitness across populations and experimental environments. We used fitness data from a common garden experiment conducted at three sites, Valencia (Spain), Norwich (United Kingdom), and Halle (Germany) (Wilczek et al., 2014). To maximize overlap between fitness and morphological data, and to remove non-genetic components of petal size variation, we predicted petal size for all 1,135 re-sequenced *A. thaliana* accessions using coefficients from a Bayesian sparse linear model ((Zhou et al., 2013), see Methods). We then compared the predicted petal size with fitness estimates in each experiment. Consistent with expectations of the selfing syndrome, genotypes carrying small-petal alleles generally performed better than those carrying large-petal alleles across most environments (**Fig. 2e and Fig. S5**).

However, the relationship between petal size and fitness varied across sites and seasons. In most cases, a quadratic model best described this relationship, with fitness peaking at relatively small petals in Norwich and Valencia, and at intermediate petal sizes in Halle. Under increased environmental stress, such as during summer and autumn in Norwich, the relationship shifted to a negative linear trend, indicating a trade-off between flower size and fitness in harsher conditions. Similarly, other datasets showed a negative association between petal size and fitness, though this trend weakened under extreme stresses (**Fig. S5**).

Overall, these results suggest that individual performance is shaped by the genetic propensity to invest in petal growth, but the strength and form of this relationship are environment dependent. Such context dependency likely generates heterogeneous selection regimes on petal size across the geographic range of *A. thaliana*, possibly explaining the lack of unidirectional selection at the species level.

Habitat suitability influences the selection regime at flower size loci.

Our analyses so far suggest that the fitness costs and benefits of investing in floral growth in *A. thaliana* depend on local environmental conditions. To examine this further, we quantified habitat suitability at the collection sites of the studied populations. We used the MaxEnt (Phillips et al., 2006) to model the species' climatic niche based on bioclimatic variables (Karger et al., 2017), altitude (Danielson & Gesch, 2011), and a comprehensive occurrence dataset (Yim et al., 2022). From this model, we extracted the habitat suitability of each studied population (**Fig. 3a, Fig. S6**) and the limiting environmental factors (**Fig. S7**).

Across populations, petal area was positively associated with habitat suitability (**Fig. 3b**). However, the relationship was triangular rather than linear: petal size variance increased toward more suitable environments, resembling the previously described relationship between flowering time and habitat suitability (Yim et al., 2022). Quantile regression confirmed this pattern, with the 0.05 and 0.95 quantile slopes differing significantly. This indicates that petal size evolution is more constrained at the climatic margins of the species' range, where environmental stress limits expansion (Yim et al., 2022), while selection on flower size is more relaxed in favourable habitats.

To test whether this triangular pattern is general across *A. thaliana phenotypes*, we assessed the relationship between habitat suitability and trait variance in multiple datasets (Atwell et al., 2010; Przybylska et al., 2023). None of the 107 traits in Atwell et al. (2010) showed significant heteroscedasticity, and only whole-plant dry mass in Przybylska et al. (2023) exhibited a similar positive triangular pattern. In our dataset, sepal length, in addition to petal area, displayed both a positive and triangular relationship with habitat suitability. Several other traits showed consistent but homoscedastic responses, including positive correlations for sepal size (this study), flowering time and leaf dry mass (Atwell et al., 2010; Przybylska et al., 2023), and negative correlations for leaf length, rosette diameter, fruit length, and leaf nitrogen content. Together, these results suggest that plant growth and life-history traits broadly align with habitat suitability, and that variance in floral organ size is particularly reduced under less favorable conditions.

To explore the evolutionary basis of this pattern, we examined how petal size alleles vary along the habitat suitability gradient. Consistent with the phenotypic trend, large-petal alleles increased in frequency in more suitable environments (**Fig. 3d**). We tested whether this could reflect variation in purifying selection intensity across the species' range. Purifying selection suppresses deleterious mutations, leading to an excess of rare derived alleles and a skewed site frequency spectrum (SFS). We computed the unfolded SFS for genes associated with petal and leaf traits to compare these effects. Low-frequency derived alleles ($0.05 < \text{frequency} < 0.1$) represented about one-fourth of SNPs in both gene sets (**Fig. 3e**). The SFS of petal genes displayed significantly higher skewness in low-suitability habitats, indicating stronger purifying selection at the environmental margin. By contrast, SFS distributions for leaf genes did not differ along the gradient.

These results indicate that genes influencing flower size are more constrained in stressful environments and experience relaxed purifying selection in favorable habitats. The accumulation of large-petal alleles in populations from suitable regions therefore likely reflects a release from the fitness penalties imposed by producing large flowers.

Discussion

Contrary to predictions from resource allocation theory, our findings indicate that petal size variation in *Arabidopsis thaliana* is not governed by a uniform trade-off with fitness. Instead, it reflects a diversity of local evolutionary pressures. Although producing large flowers can be costly, particularly in a species capable of generating up to a thousand flowers (Exposito-Alonso et al., 2018), this constraint is most pronounced in environmentally marginal habitats, where limited resources favor smaller flowers. In

contrast, in more favourable environments, such costs appear relaxed, allowing mutations that increase petal size to persist or even spread. Such spatial variation in selection likely explains the greater phenotypic variance in flower size among genotypes from climatically suitable regions.

A caveat of our study is that it does not explicitly quantify gene-by-environment (G × E) interactions, which are inherently difficult to assess across the broad environmental gradient occupied by *A. thaliana*. Our measured petal size may not fully represent the phenotype expressed under natural conditions, potentially limiting our ability to infer precise fitness relationships. Nevertheless, previous work shows that G × E effects on flower size are modest, largely influencing the magnitude rather than the direction of responses to environmental variation (Wiszniewski et al., 2022). Consequently, petal size rankings among genotypes tend to remain stable across environments, as reflected in the significant correlations between our predicted petal area and independent measurements of flower diameter at 17°C ($R = 0.35$, P-value < 0.01), flower diameter at 23°C ($R = 0.46$, P-value < 0.01) (Wiszniewski et al., 2022). Moreover, the selective signatures observed at genes controlling flower size parallel the phenotypic trends, suggesting that functional constraints on small-flower alleles intensify toward the species' climatic limits.

In addition to a relaxation of purifying selection, we detected evidence of positive selection acting on derived large-petal alleles, particularly in environmentally favorable regions. This suggests that increased petal size may confer adaptive advantages under certain conditions, or at least when environmental constraints on performance are minimal. The adaptive value of these alleles may relate to the persistence of a low but nonzero rate of cross-fertilisation in A. thaliana (Abbott & Gomes, 1989; Hoffmann et al., 2003), potentially favored during post-glacial expansions northward, following secondary contact in the Iberian Peninsula (Fulgione & Hancock, 2018). Until today, the natural variation of selfing rate (s) across a large sample of *A. thaliana* populations has been reported once (Platt et al., 2010). Consistently, we detected a positive correlation between the outcrossing rate (log(1-s)) and petal area ($R = 0.25$, P-value = 0.12) and Petal length ($R = 0.32$, P-value = 0.05) (**Fig. S8**). However, given the modest sample size (overlapping genotypes, N=39) and multiple-testing context, these relationships should be interpreted cautiously until validated by targeted analyses.

Other ecological factors may also promote larger flowers and should be explored in future studies, including historical interactions with floral enemies (Galen, 1999), protection against microbial and viral infections (Vannette, 2020), and the role of flower thermogenesis in pollen germination (van der Kooi et al., 2019). Pleiotropy and linked selection could further contribute to the emergence of large-flower alleles, especially as some petal size–associated variants lie within genes that influence other organs or developmental pathways (Hanemian et al., 2020). While our genetic correlation analyses indicate that such effects are limited, functional validation of causal mutations will be necessary to strictly disentangle developmental coupling from direct adaptive responses.

Overall, this study reveals striking geographic variation in investment in reproductive structures, particularly flower size, across the range of *A. thaliana*. Classic theory predicts that selfing lineages experience strong selection for reduced floral display, leading to the fixation of small-flower alleles due to their lower resource costs. Our results support this expectation only under environmentally stressful

conditions, where resource limitation likely intensifies selection for smaller flowers. In contrast, in favourable environments, relaxed purifying selection reduces these constraints, enabling the persistence or even evolution of larger-flower variants. Together, these findings illustrate how habitat suitability and resource availability modulate the balance between selection and constraint, shaping the trajectory of phenotypic evolution in selfing species.

Materials and Methods

Biological materials

For this study, we subsampled 407 *Arabidopsis thaliana* accessions from (Przybylska et al. 2023 data), originally selected for their extensive geographic and genetic coverage (**Fig. S1, Dataset S1**). All accessions were vernalized for 4 weeks in the dark at 5°C in October 2022. Seedlings were then transferred to soil in individual pots and grown in a glasshouse in semi-controlled conditions in Montpellier, CEFE CNRS, France (maintained around 20°C day and 15°C night with natural light). Accessions were randomly distributed on three tables daily rotated to homogenize the effect of spatial heterogeneity within the glasshouse across the genotypes. We used a single replicate per accession which allowed to maximize the genetic diversity coverage in a restricted experimental area, while genetic relatedness between accessions provided partial genetic replicates. Variance analysis within and between *Arabidopsis thaliana* regional populations (Alonso-Blanco et al., 2016) revealed high heritability for all investigated traits (**Dataset S2**), indicating strong phenotypic similarity among groups. Plants were surveyed daily, and the date of first flower opening, flower collection and plant death were recorded. Three flowers were collected between the 10th and the 25th flower of the main stem from the primary inflorescence. Flowers were collected and preserved in 70% ethanol in individual Eppendorf placed in the dark at room temperature.

Morphological measurements

Flowers were stained with 0.02% Toluidine Blue for 10 minutes before been dissected and placed in between microscope slides and cover slips with a drop of glycerol (40% v/v). Slides were scanned at a 4800 dpi resolution with an Epson perfection v850 pro® scanner. Leaf, petal, sepal, and short and long stamen sizes (length, width, and area when relevant) were measured using Fiji version 2.35 (Schindelin et al., 2012). The blue channel of the images was isolated, and contrast-enhanced with Otsu's local contrast method. Pixels that belong to the organs were counted to estimate organ sizes. Measures were calibrated using the dots per inch (dpi) of the images. The ovule number per flower was counted manually using fluorescence microscopy (Leica DMI4000). Flowering time was calculated as the duration in days from seed sowing to the opening of the first flower. To estimate the genotype leaf sizes, while limiting environmental effects, the fourth leaf of every individual was collected at maturity and immediately scanned at 800dpi resolution. We computed an unbiased predictor of each trait by extracting the residuals of a linear model containing the relevant confounding factors (petal rank along the stem, stem number and a variable describing the spatial distribution in the tables). All traits were normally distributed and did not necessitate transformation (**Fig. S2**). We used our phenotypes in conjunction with published datasets that shared a significant amount of accessions. In particular, we

used the 107 phenotypes gathered by (Atwell et al., 2010), containing phenology, microbial resistance, plant morphology and ions content values, and the Aradiv database from (Przybylska et al., 2023) that contains plant functional strategies related traits. We also used datasets reporting fitness in terms of seed production where a majority of our accession set was shared, such as in (Exposito-Alonso et al., 2018; Wilczek et al., 2014), and compared our measurement of flower organ size with other available datasets (Li et al., 2020; Wiszniewski et al., 2022).

Statistical analyses

All statistical analyses were performed with R (version 4.3.2). We performed principal component analyses with *FactoMineR*. Testing for isometry versus allometry was achieved with standard major axis regression, which assumes an error component from both tested variables, with the package *smatr* (Warton et al., 2006). Skewness and deviation from the average were statistically tested in R using a skewness test and a Student's t-test, respectively.

Genome-Wide Association Study (GWAS)

We performed GWAS by testing the association between the above measured organ sizes with publicly available variant calling format (VCF) data from the 1001 genomes project webpage (https://1001genomes.org/data/GMI-MPI/releases/v3.1/) (Alonso-Blanco et al., 2016). Although single replicate per accession may reduce detection power by introducing environmental noise, the high heritability of the measured traits and population structure signals associated with phenotypic variation suggest that increasing replicates per accession would have minimal impact on genetic association power (Kang et al., 2008). Moreover, our measurement of petal size correlated significantly with independent flower diameter measurements, i.e. the ranking from small to large flower genotypes remains consistent in between treatments (Wiszniewski et al., 2022) and across experiments (Spearman's Rank Correlation Coefficient ~ 0.46). Only single nucleotide polymorphisms (SNPs) were kept in the analyses, short indels and non-ACTGN were filtered out, as well as SNPs with a minor allele frequency below 5% using PLINK (Purcell et al. 2007). A total of 1,899,962 SNPs passed the filters. GWAS were performed using the GEMMA software v0.98.1 (Zhou & Stephens, 2012), which considers the population structure in a relatedness matrix. The proportion of phenotypic variance explained by the genotypes (PVE, the SNP-based heritability) was estimated for each trait with the bslmm function (Bayesian sparse linear mixed model). The P-values distribution of all GWAS followed a uniform distribution and did not necessitate transformation for false discovery rate. Specific quantitative trait loci (QTL) associated with the traits were located in two ways. First we sought major QTL by a genome-wide screening of the association significance, accounting for multiple testing with Bonferroni corrected significance threshold. Second, to consider the potential polygenic architecture of the studied traits, we relaxed the threshold of significance. As there is no consensus on a threshold to detect multiple minor effect variants in a polygenic architecture (also known as multiple disease phenotype), reports of P-values detecting relevant genes range from 10-7 to 10-4 (Burton et al., 2007; Chen et al., 2021), we used an empirical approach and set the threshold at a P-value where we detected the FLOWERING LOCUS C (FLC) among the flowering time GWAS association peaks (P-value of 5.2*10-5). FLC, known

as a key regulator of flowering time, is commonly detected in GWAS in A. thaliana, and its variation accounts for more than half the trait's variation (Sasaki et al., 2018). To eliminate genetically linked variants, we filtered out significant SNPs by considering their genomic position (https://www.arabidopsis.org), statistical significance, and predicted function using SnpEff (Cingolani MMB, 2012). This process resulted in retaining one focal polymorphism per QTL for subsequent analyses (**Dataset S**4). SNP effect sizes were estimated from the bslmm outputs, as recommended by Zhou et al. (2013) by the formula α + β.γ, with α the small additive effect of each SNP, β the additional effect of some large effect SNPs and γ its posterior inclusion probability estimate. Estimation of the coefficients were further used to predict traits values for all 1,135 sequenced genotypes of the 1001 genome project using the predict function from GEMMA (Karasov et al., 2024; Lo et al., 2024).

Population Genetics Analysis

We assessed the regime of selection acting on leaf and flower size QTLs across the species distribution with two metrics. We first computed the πN/πS ratio for all *A. thaliana* annotated genes. We calculated the ratio between the genes' nucleotide diversity in non-synonymous sites (πN – defined as the 0-fold degenerate sites) and the nucleotide diversity in synonymous sites (πS – defined as 4-fold degenerate sites). The rationale is that πS gives a measure of neutral genetic variation and serve as a reference for πN which measures the amount of genetic variation at sites under natural selection control. Thus, a value of πN/πS below 1 is indicative of purifying selection, and a value above 1 is indicative of diverging selection. We used the genome-wide distribution as a reference: individual genes' πN/πS included in the 0.05 to 0.95 quantile do not deviate from the general purifying trend, Student test and skew test were performed on R to evaluate a set of genes πN/πS deviation from average and normality. Full sequences of gene coding regions and general feature format files were downloaded from Phytozome 13 (version Athaliana_447_TAIR10). Together with the VCF file, they were used to generate the diversity of gene sequences observed among *A. thaliana* genotypes. The R package *RGenetics* was used to discriminate 4-fold from 0-fold degenerate sites, and the *nuc.div* function from *pegas* was used to compute the nucleotide diversity. More details about the methods and scripts used are available at (https://github.com/kevinfrsartori/Whole_Genome_PiNPiS). Second, the extended haplotype homozygosity (EHH) was computed genome-wide to test whether a particular derived variant is subjected to ongoing or partial sweep in populations (Sabeti et al., 2002). When advantageous mutations are selected, the frequency of genetically linked variants also increases, extending the haplotype length around the selected site. This signal perdures until recombination events erase it. EHH measure the decay of a haplotype length upstream and downstream of the focal SNP, which thus depends on the selection's history and strength. Since the patterns of EHH might also be affected by the species' population structure and demographic history (Voight et al., 2006), we compute this estimate genome-wide to assess significance. The R package *rehh* (Gautier and Vitalis Bioinformatics, 2012) was used to compute the EHH's test statistic (iHS), defined as the log ratio of the integral of the EHH decay (iHH) around the derived allele divided by the iHH around the ancestral allele. This statistic was then compared to the distribution of iHS for all SNPs of similar allele frequency, and a P-value was reported. The iHS computation relies on the polarization of the SNP dataset, i.e. the ancestral state of

each SNP. To differentiate between the derived versus ancestral allele, we used a maximum likelihood method using the focal species allele frequencies together with related species allelic information. We first identified an ortholog gene in *Arabidopsis lyrata* and *Arabidopsis halleri* for most *Arabidopsis thaliana* genes by using OrthoFinder (Emms & Kelly, 2015). We kept all *A. thaliana* genes having at least one ortholog and kept only one ortholog per outgroup species based on the genes' distance tree, resulting in a list of 25411 genes. We extracted a sequence covering the coding region plus the cis- and trans-regulatory regions for all trio of genes, and performed a multiple alignment using MAFFT (Katoh et al., 2002). We extracted the allelic information of the outgroup species with custom R scripts and ran the program est-sfs (Keightley & Jackson, 2018). It is important to note that the computation of iHS was limited to the presence of ortholog genes, resulting in filtering out few SNP per trait lists (**Dataset S4**).

Habitat suitability

The *A. thaliana* niche was modelled using the MaxEnt software (Phillips et al., 2006) by reproducing the method from (Yim et al., 2022). Climatic data were downloaded from CHELSA v2.1 (https://chelsa-climate.org) at a resolution of 30 arc seconds (Karger et al., 2017). The species distribution modelling was performed with the following variable panel: the isothermality (ratio of diurnal variation to annual variation in temperatures, °C), the mean of daily minimum air temperature of the coldest month (°C), the annual range of air temperature (°C), mean daily mean air temperatures of the wettest quarter (°C), mean daily mean air temperatures of the warmest quarter (°C), precipitation seasonality (standard deviation of the monthly precipitation estimates, kg m-2), mean monthly precipitation amount of the wettest quarter (kg m-2 month-1), and mean monthly precipitation amount of the driest quarter (kg m-2 month-1). Altitude data at a resolution of 30 arc seconds come from the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) (Danielson and Gesch 2011) and were downloaded from the USGS website. We used the extensive species occurrence data (N=672) provided by (Yim et al., 2022), which covers the distribution of the species in an evenly spaced manner (one sample per km2) to account for sampling bias. From the MaxEnt model, we estimated the habitat suitability (HS) and the limiting factor (LF) of all *A. thaliana* accessions' collecting sites (**Fig. S6 S7**). The HS metric scales from zero to one, while LF refers to the climatic variable that has the most significant impact on HS when changed while maintaining all others variable constant.. The relationships between traits and HS were tested using linear regressions. The apparent heteroscedasticity of several trait-HS relationships was tested by comparing the 0.95 and 0.05 quantile regressions (Peng, 2017) using the R package *quantreg*. Linear regressions were plotted when significant, and quantile regressions were plotted when the test for slope difference (Anova) was significant. All statistical analyses were performed with R (R Core Team, 2020). In order to test for selection regime shift along habitat suitability, we divided the *A. thaliana* accession set into ten equally sized sub-populations. To avoid the influence of over-sampling in some regions of the species distribution, we first pruned the accession list based on genetic relatedness. We used the PLINK software to estimate the number of accessions passing a filtering threshold of genetic relatedness ranging from r=0.1 to r tends to 1. We found that the number of related accessions stops dropping after 0.95 and remains constant while r tends to 1 (**Fig. S9**). We thus excluded 87 highly

459 correlated accessions. HS ranges were obtained by dividing the HS distribution of unrelated accessions
460 into ten quantiles, where each quantile contains about 100 accessions.

461 ***Selection regime along habitat suitability***

462 The HS ranges obtained by splitting the A. thaliana accession into 10 populations by habitat suitability
463 did not have a geographic and historic meaning for the species, which prevents from computing most
464 classical population genetic metrics. We tested for changes in selection pressure by relying on allele
465 frequency only. We used the sequencing dataset presented above and reduced the minimum allele
466 frequency threshold to a minimum allele count of 1. Only the SNPs having ancestry information located
467 in the gene body were kept, and we computed the unfolded site frequency spectrum for leaf and flower
468 organs associated genes by reporting the allele counts for ranges of allele frequency of 0.05.

469 **Acknowledgments**

470 We thank the Sicard, and Rosa groups for discussions and comments on the article. We thank
471 Alexander Platt for valuable discussions and method sharing. The computation and data handling were
472 provided by the Swedish National Infrastructure for Computing (SNIC) at Uppmax, partially funded by
473 the Swedish Research Council through grant agreement no 2018-05973. This work was supported by
474 the Swedish Research Council (grant 2018-04214) and the Novo Nordisk Foundation
475 (NNF22OC0079830) to AS and JRL was supported by NIH award R35GM138300.

476 .

477

478 # References

479

480

481

Figures and Tables

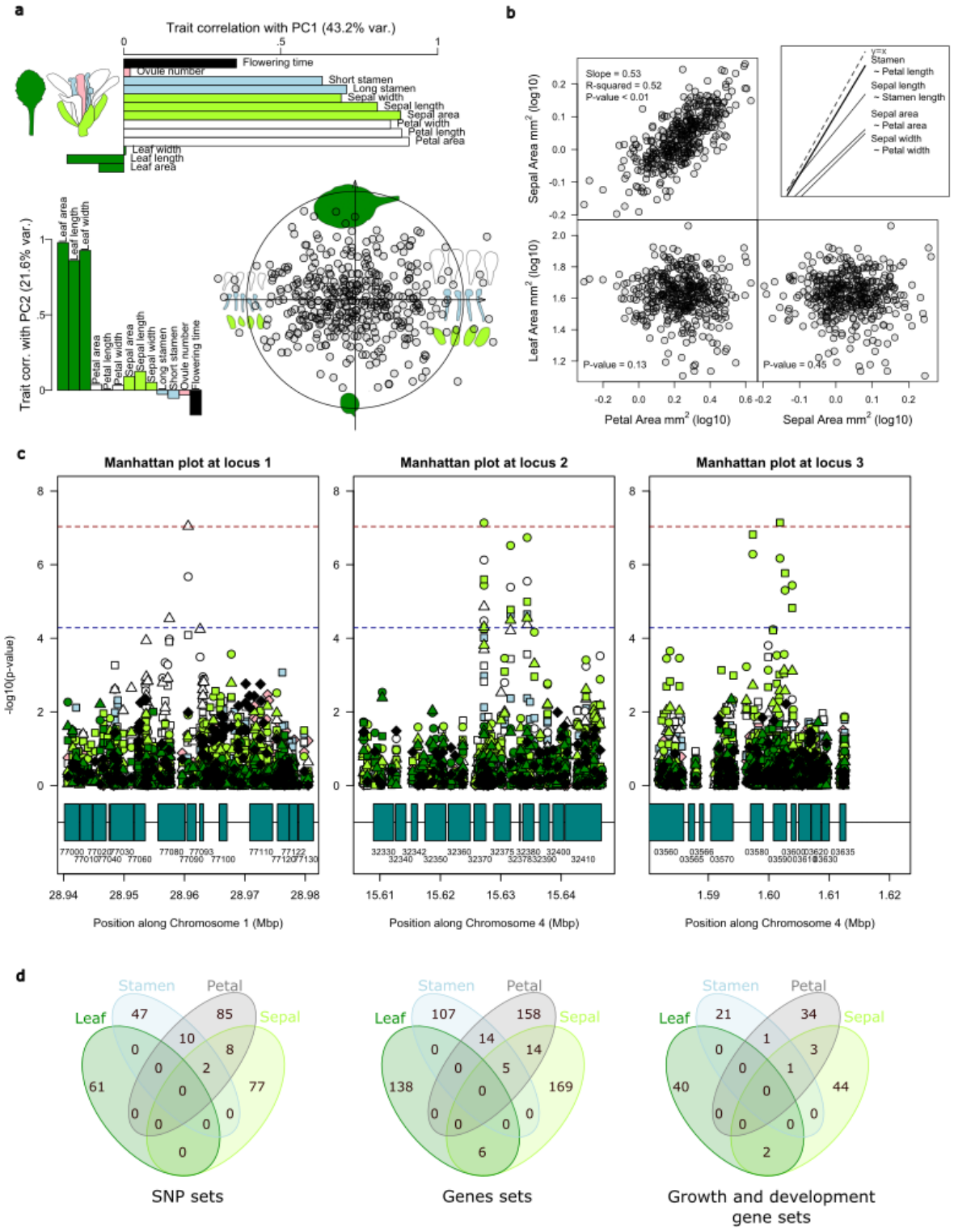

Figure 1: Extensive flower size variation in *A. thaliana* results from low pleiotropic mutations.

a) A Principal component analysis (PCA) was performed on flower organs and leaf size excluding flowering time that was correlated post hoc and represented as a supplementary variable. Bar plots show trait correlations with principal components (PC), with colours indicating different organs as shown on the top right scheme: dark green, leaf; light green, sepal; white, petal; blue, stamen; pink, ovules. The biplot displays accession coordinates along the first two principal components (PC), with organ silhouettes representing extreme organ size values at the same scale; **b)** Pairwise relationships between organ sizes. Coefficients were obtained from standard major axis linear regressions. The top right plot illustrates the slope test for organ isometry (null hypothesis: slope = 1). **c)** Manhattan plots displaying the genomic regions surrounding the three significant associations peaks (based on Bonferroni threshold) detected for all studied traits. Dot colours indicate associations with the size of different organs as depicted in a), while shapes correspond to the parameter analyzed (triangles indicate association with area, circles with width or short, and squares with length or long). The Red dashed line represents the Bonferroni threshold, while the blue dashed line represents the polygenic threshold. Blue rectangles materialize the annotated gene code numbers (following “ATxG” in TAIR nomenclature, with *x* being the chromosome number). **d)** Venn diagrams illustrating the lack of overlap between SNPs, genes and candidate genes associated with leaf and flower organs’ size variation.

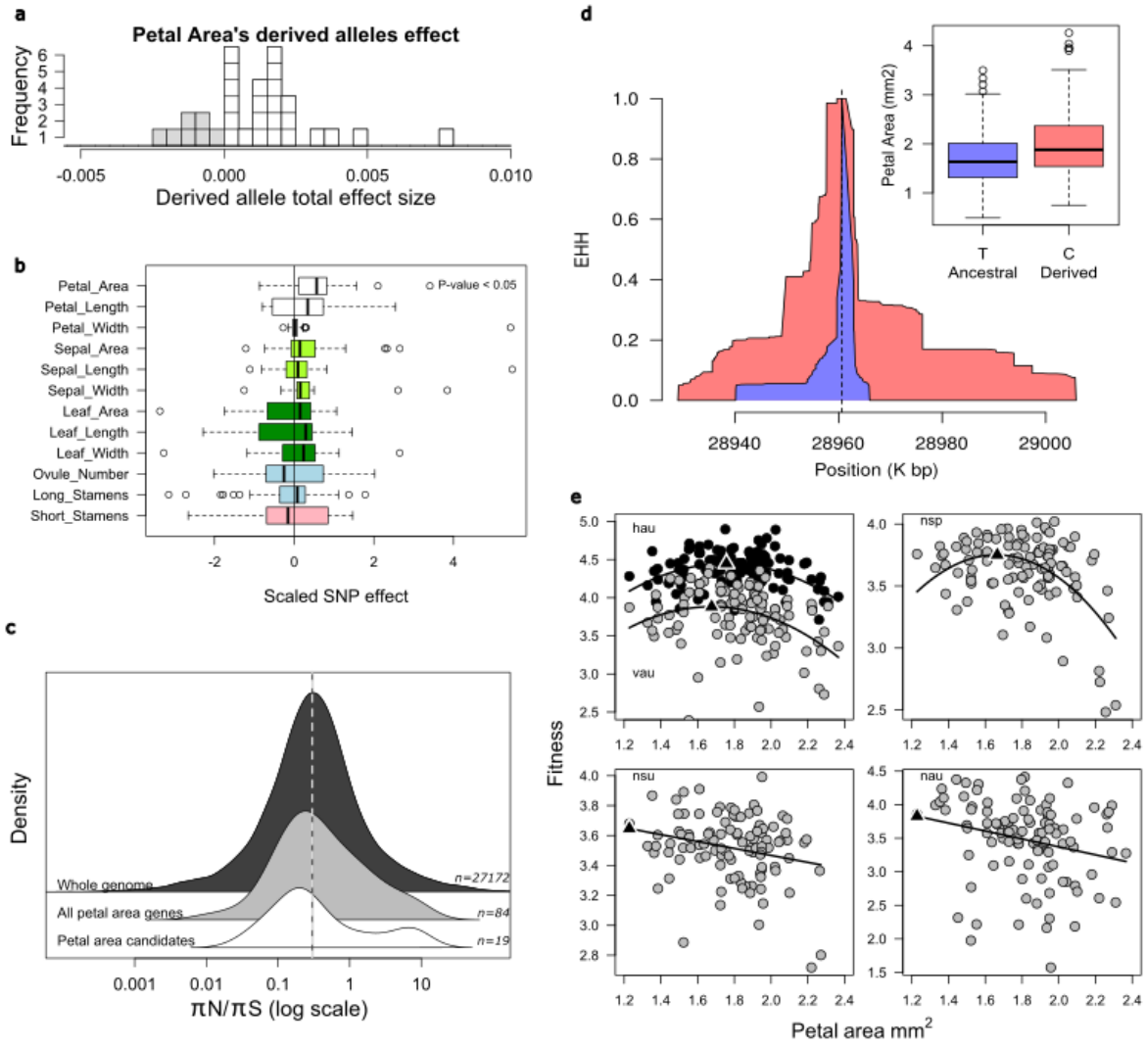

Figure 2: A variable flower size–seed production trade-off explains a mixed selection pattern on flower size in *A. thaliana*.

a) Distribution of derived alleles' total effect size from each peak of association identified in GWAS for petal area. Each square represents a SNP's derived allele, white and grey indicating a positive and negative effect, respectively. **b)** Distribution of derived alleles' total effect size from each peak of association identified in GWAS for each trait investigated. A p-value is given for trait distributions significantly different from 0 (vertical line). **c)** πN/πS distribution across all *Arabidopsis thaliana* genes, among the petal area genes and a priori candidate genes identified in the GWAS for petal area. The dashed line represents the whole genome average. **d)** Extended haplotype homozygosity around FLM lead SNP position. The red and blue integrals represent the degree of haplotype identity shared by individuals carrying the C (derived) and T (ancestral) alleles, respectively, as a function of distance to the lead SNP (dashed line). **e)** Relationship between the genetic variance of petal area and seed

520 production across different experimental fields; hau, Halle autumn; vau, Valencia autumn; nsp,
521 Norwich spring; nsu, Norwich summer; nau, and Norwich autumn. Triangles represent the
522 modelled petal size that maximizes seed production, black lines represent the best fit (anova)
523 between a linear and a quadratic regressions, and blue shadows represent the confidence
524 intervals of the regressions.

525

526

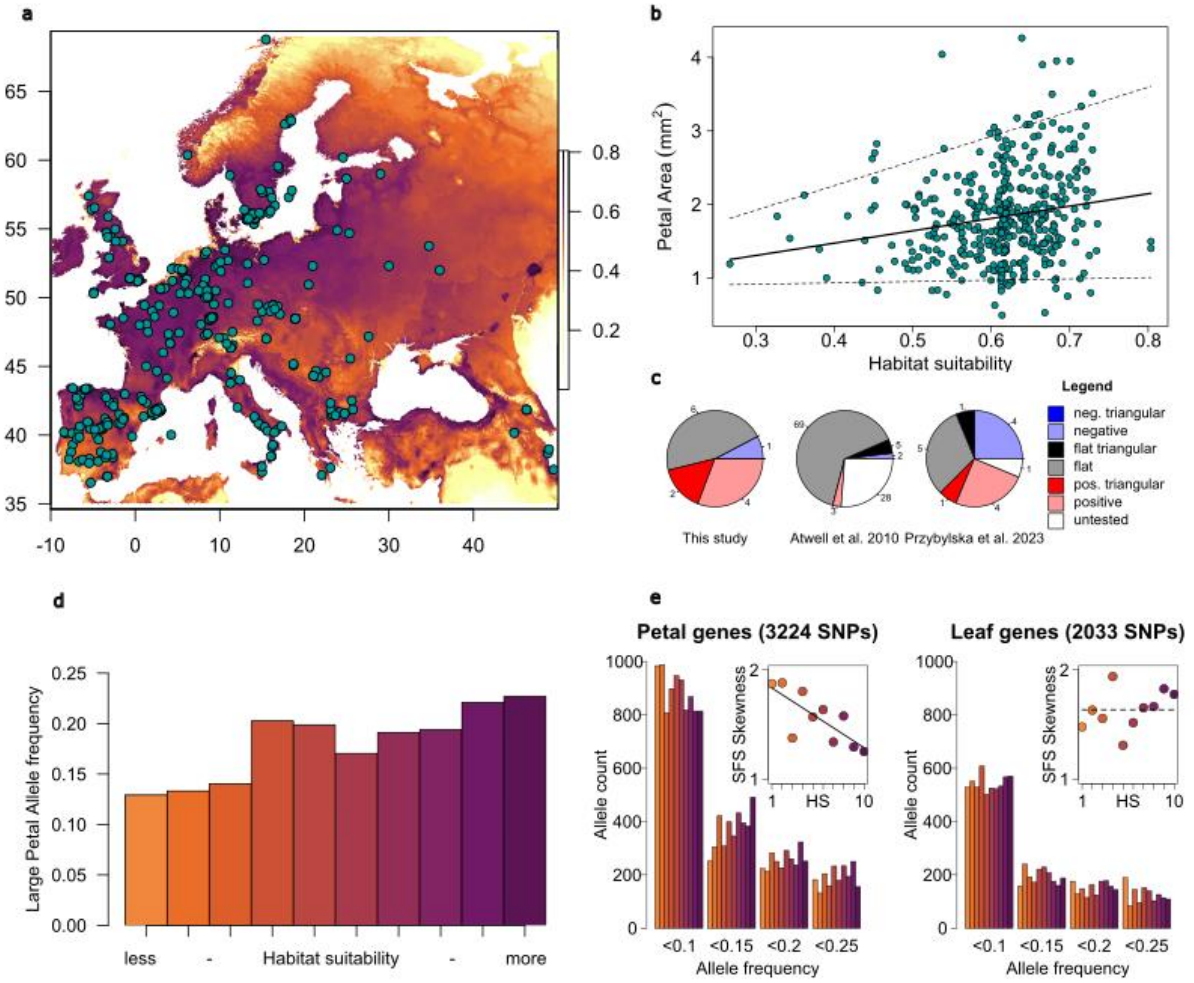

Figure 3: Habitat suitability shapes patterns of phenotype diversification across the geographical range of *A. thaliana*.

a) Map of the studied European accessions. The orange to purple scale represents the modelled habitat suitability from less to more suitable, and the blue dots represent accessions' collecting sites. **b)** Correlation between petal area and habitat suitability. Lines represent significant regressions: plain line, linear model; dashed lines, 0.05 and 0.95 quantile regressions. **c)** Large petal alleles' frequency along habitat suitability clusters. **d)** Unfolded site frequency spectrum (SFS) among SNPs located in petal and leaf gene sets. The colours of the bars correspond to the habitat suitability classes. Inserts represent the relationship between the SFS's shape and habitat suitability, where a plain line depicts significant and dashed line insignificant correlation.