# Genetic Architecture of Flowering Time Differs Between Populations With Contrasting Demographic and Selective Histories

Célia Neto [ID][‡] and Angela Hancock [ID]*

Molecular Basis of Adaptation Research Group, Max Planck Institute for Plant Breeding Research, Cologne, Germany
‡Present address: Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark
**Corresponding author:** E-mail: hancock@mpipz.mpg.de.
**Associate editor**: Dr. Michael Purugganan

## Abstract

Understanding the evolutionary factors that impact the genetic architecture of traits is a central goal of evolutionary genetics. Here, we investigate how quantitative trait variation accumulated over time in populations that colonized a novel environment. We compare the genetic architecture of flowering time in *Arabidopsis* populations from the drought-prone Cape Verde Islands and their closest outgroup population from North Africa. We find that trait polygenicity is severely reduced in the island populations compared to the continental North African population. Further, trait architectures and reconstructed allelic histories best fit a model of strong directional selection in the islands in accord with a Fisher–Orr adaptive walk. Consistent with this, we find that large-effect variants that disrupt major flowering time genes (*FRI* and *FLC*) arose first, followed by smaller effect variants, including *ATX2* L125F, which is associated with a 4-day reduction in flowering time. The most recently arising flowering time-associated loci are not known to be directly involved in flowering time, consistent with an omnigenic signature developing as the population approaches its trait optimum. Surprisingly, we find no effect in the natural population of *EDI-Cvi-0* (*CRY2* V367M), an allele for which an effect was previously validated by introgression into a Eurasian line. Instead, our results suggest the previously observed effect of the *EDI-Cvi-0* allele on flowering time likely depends on genetic background, due to an epistatic interaction. Altogether, our results provide an empirical example of the effects demographic history and selection has on trait architecture.

*Key words*: directional selection, flowering time, complex traits, Fisher–Orr geometric model, adaptive walk.

## Introduction

Understanding how phenotypic variation is generated and maintained within populations and how this variation is encoded in the genome are central goals of evolutionary genetics (Barton and Keightley 2002; Lee et al. 2014; Josephs et al. 2017). Since Darwin and Wallace and the origin of evolutionary theory, there has been a debate over the relative importance of jumps and leaps versus gradual change in trait evolution (Wallace 1855; Darwin 1859; Gould and Eldredge 1972; Provine 2001). Developing a quantitative understanding of trait architectures and what shapes them is important in both fundamental and applied research. Specific parameters of interest include the number of variants that contribute to observed variation in a trait, the frequencies of those variants in the natural population, and their distributions of effect sizes. Estimates of these parameters in natural populations are needed to parameterize theoretical models, predict evolutionary outcomes under future scenarios, and design powerful trait-mapping studies. However, our understanding of the range of genetic architectures in nature and the factors that influence them is still remarkably limited.

Theory predicts that several parameters are likely to shape genetic architectures of quantitative traits in natural populations. These include the amount of heritable genetic variation available within a population, which is related to the long-term effective population size ($N_e$), or the inverse of the time depth of the population, as well as the selective force and strength of selection acting on the trait (Gillespie 1983b; Orr 1998; Silander et al. 2007; Gomulkiewicz and Houle 2009; Gomulkiewicz et al. 2010; Barton et al. 2017; Sella and Barton 2019; Charlesworth et al. 2022). In large-$N_e$ populations, if the trait of interest is evolving under stabilizing selection, new mutations with substantial effects on fitness-associated traits are likely to be removed rapidly by purifying selection. Variants found segregating in these large-$N_e$ populations are therefore expected to be at low frequency, and any specific mutation should have only a minor impact on a trait. Conversely, in a population that experiences strong directional selection, beneficial large-effect variants may provide the first steps toward adaptation (Orr 1998, 2005).

Fisher's "infinitesimal model" assumes a population of infinite size evolving under stabilizing selection

**Open Access**

(Fisher 1918). As new mutations arise, they are expected to result in minute effects across the genome that pull individuals away from the fitness optimum and contribute to trait variation within the population. Fisher extended his infinitesimal model to include adaptation in a "geometric model" (Fisher 1930). Under this model, a population far from a high-dimensional phenotypic optimum moves closer to it using random mutations that arise over time. Fisher assumed that universal pleiotropy constrains evolution, providing further rationale for the idea that many variants of small phenotypic effect should underlie trait variation in the population (Fisher 1930). Recent work on human disease trait evolution, which is based on Fisher's infinitesimal and geometric models, shows support for these models. These studies examine the impact of variation arising in the context of a trait evolving under stabilizing selection (as is the case for many human disease traits) (Simons et al. 2018; Sella and Barton 2019) or after a weak or moderate shift in a selective pressure and thus a phenotype (Hayward and Sella 2022; Simons et al. 2022). However, the infinitesimal model is not sufficient to explain patterns that arise under very strong directional selection (Barton et al. 2017).

In the case where a very large environmental change occurs and available genetic variation cannot produce an adequate phenotypic shift, new large-effect loci are expected to play a prominent role in the adaptive process (Gillespie 1983b; Matuszewski et al. 2014; Orr 1998, 2002). Building on Fisher's model and subsequent theoretical advances (Kimura 1983; Gillespie 1983a; Kauffman and Levin 1987), Orr formulated a model of the genetics of adaptation that approximates theoretical and empirical data (Orr 1998). The resulting Fisher–Orr model asserts that, after a sudden change in the environment (e.g., following a major environmental perturbation or after colonization of a new habitat), a mutation-limited population will move toward the new fitness optimum through an "adaptive walk". Under this model, the first steps of adaptation are likely to occur through few a large-effect mutations that overcome genetic drift to increase in frequency in the population (Orr 1998, 2002). Later in the adaptive walk, smaller effect mutations fine-tune phenotypic variation (Orr 2002). Overall, effect sizes are expected to follow an exponential distribution, with a few large impact variants and many small impact variants. This scenario can be contrasted with the expectation under stasis, where stabilizing selection mainly acts to preserve trait values that maximize fitness against newly arising deleterious variation. Under these conditions, the distribution of effect sizes is expected to be uniform, although it has been suggested that even under neutrality an exponential or similar distribution might be expected (Robertson 1967; Cotto and Day 2023). Experimental work on bacteria and yeast (Barrick et al. 2009; Frenkel et al. 2014; Good et al. 2017) has provided some empirical insights on how populations adapt over time as populations evolve. However, real populations may be more complex than these models assume. For example, real populations may be spatially structured and/or subject to spatial and temporal variation in selection pressures (Rausher and Delph 2015; Dittmar et al. 2016; Connallon and Hodgins 2021).

While theory provides expectations about how demographic and selective histories should impact trait architecture, natural systems that allow us to make direct comparisons are rare. For this, we need to study populations with known, contrasting histories. Here, we compare the genetic architecture of flowering time in *Arabidopsis* populations that colonized the Cape Verde Islands (CVI) with that of their closest outgroup population from Morocco.

The colonization history of *Arabidopsis thaliana* in the CVI is well-defined (Fulgione et al. 2022), providing an effective case for investigating how population history and selection impact trait architecture. The CVI lies at the geographic and climatic edge of the species distribution. Natural populations of *A. thaliana* populate two CVI: Santo Antão and Fogo, where they are restricted to altitudes greater than 900 m. The islands were colonized approximately 4–5 kya by long-range dispersal from a North African progenitor (fig. 1), best represented today by the Moroccan population (Fulgione et al. 2022). Consistent with colonization through an extreme bottleneck, diversity is strongly reduced in CVI relative to Morocco (Durvasula et al. 2017; Fulgione et al. 2022). Estimates of long-term $N_e$ based on diversity ($\theta_W$) and estimated mutation rate (Ossowski et al. 2010; Exposito-Alonso et al. 2018; Wang et al. 2019; Weng et al. 2019; Fulgione et al. 2022) are approximately 60 to 90 times higher in the Moroccan population compared to Santo Antão and Fogo, respectively ($N_{e\ Santo\ Antão} =$ 2–4 K, $N_{e\ Fogo} =$ 3–5 K, $N_{e\ Morocco} =$ 176–323 K, for a mutation rate between $4.3 \times 10^{-9}$ and $7.9 \times 10^{-9}$). The near-complete bottlenecks during colonization eliminated 99.9% of the shared variation with the continent, producing phylogenetically distinct variation-depauperate populations in each of the two islands (Fulgione et al. 2022) in which it is straightforward to assign ancestral states to variants. This case provides the possibility to directly compare trait evolution in the long-established, high-$N_e$ Moroccan population, with a population coalescence time of approximately 1 mya (Durvasula et al. 2017), and two distinct recently established Cape Verde island populations, where variation is expected to have arisen de novo since colonization (Fulgione et al. 2022). This situation in some ways parallels experimental evolution studies, where it is possible to investigate how mutations accumulated over time in the population.

Reproductive timing is often a crucial element of an organism's life cycle and a primary fitness component in nature (Stearns 1992; Fabian and Flatt 2012). In plants, flowering time is an excellent model trait for connecting ecological and evolutionary factors with their impacts on fitness. Flowering too early—before an individual has had time to accumulate sufficient resources—may limit reproductive capacity. Conversely, flowering too late can result in low or zero fitness in drought-prone environments
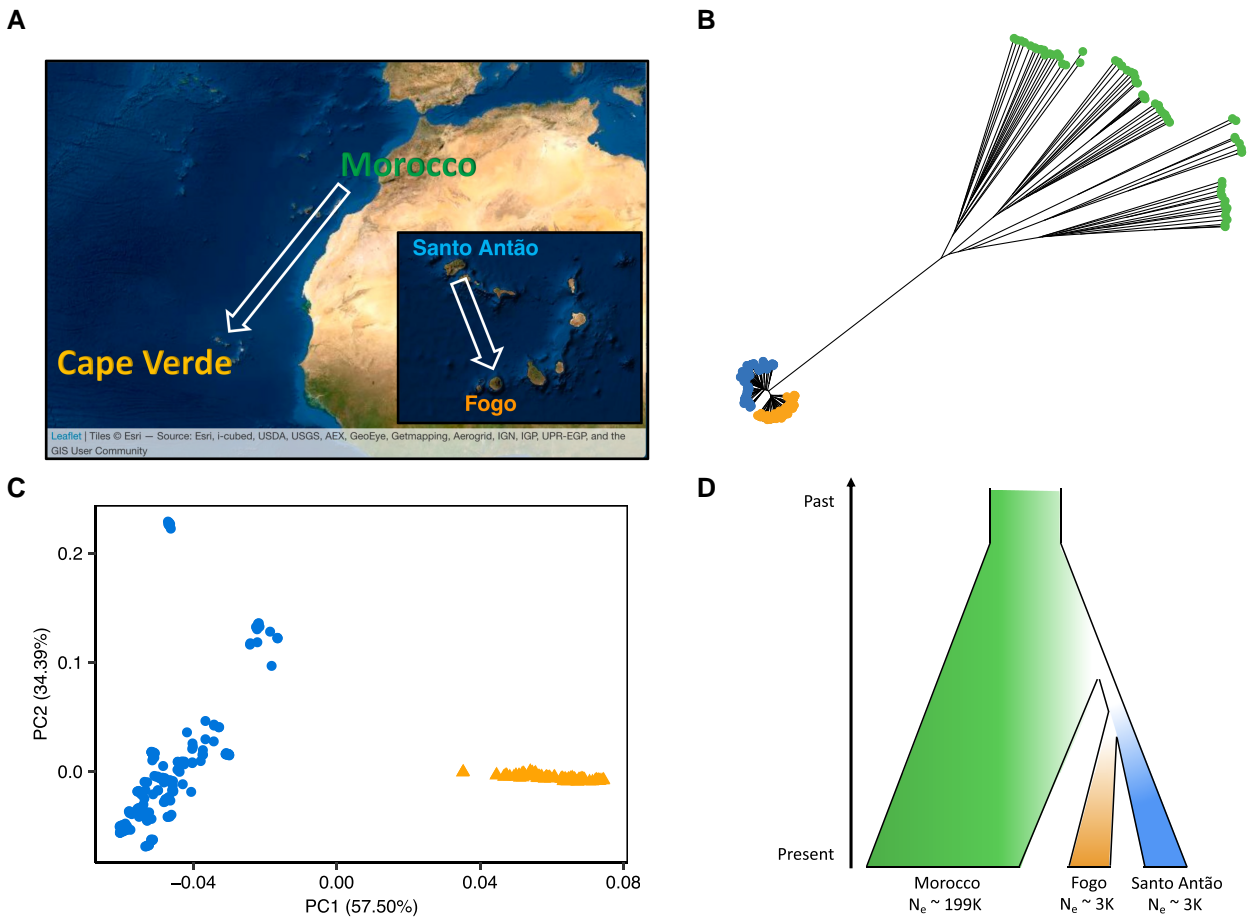
**Fig. 1.** Geographical location and population history of CVI Arabidopsis. (A) Arabidopsis colonized Cape Verde 4–5 kya from North Africa. (B) Neighbor-joining tree showing the relationship between Moroccan and Cape Verdean individuals. (C) Principal component analysis (PCA) showing the differentiation between Santo Antão (circles) and Fogo individuals (triangles). (D) Schematic of Cape Verde Arabidopsis history.

(Ludlow 1989; McKay et al. 2003). Several studies have shown that flowering time is an important component of reproductive success (Hall and Willis 2006; Franks et al. 2007; Korves et al. 2007; Li et al. 2010; Anderson et al. 2011; Ågren and Schemske 2012; Dittmar et al. 2014; Ågren et al. 2017), a finding that is further supported by observed clinal patterns (Ducousso et al. 1996; Hurme et al. 2000; Olsson and Ågren 2002; Caicedo et al. 2004; Stinchcombe et al. 2004).

We previously showed that flowering time is reduced in CVI Arabidopsis populations, which increases reproductive success through a drought escape mechanism (Fulgione et al. 2022). We identified two large impact variants responsible for the convergent reduction in flowering time between the two CVI, Santo Antão and Fogo. These variants disrupted two interacting genes (FRI in one island and FLC in the other), reducing the flowering time by 27–31 days in each case (Fulgione et al. 2022). However, since flowering time is a complex trait, its basis is expected to be polygenic, with contributions from many loci genome-wide (Mouradov et al. 2002; Salomé et al. 2011; Andrés and Coupland 2012; Bouché et al. 2016; Zan and

Carlborg 2018, 2019). In this mutation-limited population facing an abrupt environmental change, we expect adaptation to follow an adaptive walk toward the new optimum. According to theory, large-effect mutations should then arise early in the walk, followed by smaller ones later, with the total effect size distribution roughly fitting a negative exponential (Orr 1998).

Here, we examine the full genetic architecture of the time to first flowering in Cape Verde and compare it to the architecture in the Moroccan outgroup population. We find that polygenicity is severely reduced in the colonizing populations, consistent with the recent coalescent times of these populations. We identify loci associated with the reduction in the time to first flowering, hereafter flowering time, on the islands and show that these have smaller effects and arose more recently than the two large-effect loss-of-function mutations previously identified in CVI (FRI K232X and FLC R3X) (Fulgione et al. 2022). Overall, this approach allows us to examine how polygenic architecture built up over time in an ecologically relevant trait and provides an empirical example of the effects of demographic history and selection on trait architecture.

# Results

## Genome-Wide Association Studies Peaks Contain Loci Implicated in Core Flowering Time Pathways as Well as Peripheral Pathways

We propagated lines originally sampled from the two CVI and Morocco (fig. 2a) together in a growth chamber programed to simulate CVI field conditions, based on data we collected using data loggers placed in Cape Verde. These included light intensity, basic soil characteristics, daily precipitation, and hourly temperature and humidity. We measured the time it took for plants to bolt, hereafter referred to as "flowering time". On average, flowering time was faster in Cape Verde compared to Morocco and inversely correlated with seed production in a simulated CVI environment (fig. 2b–c) (Fulgione et al. 2022). We calculated broad-sense heritability based on repeatability across replicates in an analysis of variance framework (line effect), as well as narrow-sense heritability based on the proportion of trait variance explained by relatedness (PVE) (Zhou et al. 2013). Heritabilities were high (table 1): The line effect ranged from 51% to 85% in a model that controlled for block effects and the PVE based on the trait correlation with the relatedness matrix was greater than 95%, implying that a large proportion of the variation in flowering time is genetic.

To identify the genetic loci underlying flowering time variation, we conducted genome-wide association studies (GWAS) separately for each population using a linear mixed model (LMM) approach that controls for relatedness using a kinship matrix (Zhou and Stephens 2012) (supplementary fig. S1, Supplementary Material online). Since we already had clear functional evidence that two variants, FRI K232X in Santo Antão and FLC R3X in Fogo, strongly affect flowering time in CVI (Fulgione et al. 2022), we accounted for these in subsequent analyses. In particular, since FRI K232X segregates in Santo Antão, we included this variant as a covariate in the LMM. To reduce redundancy resulting from local linkage disequilibrium (LD) and to increase the power to detect associations across haplotypes, we used an approach that aggregates association signals based on LD between markers within a genomic region (Fariello et al. 2017; Bonhomme et al. 2019).

In this method, P-values from GWAS are integrated across haplotypes along genomic regions based on LD (Mercier and Daudin 2001; Fariello et al. 2017; Bonhomme et al. 2019). These are converted to "local scores", such that strongly associated regions have the most positive scores. In effect, this method identifies variants that represent the signal in larger haplotypes. The local score approach is in some ways analogous to window-based identification of candidate loci, but it has been shown to have higher power than these approaches (Fariello et al. 2017).

In Santo Antão, we identified five significantly associated regions (fig. 3), containing variants in seven genic candidate loci. Among these were several genes with known direct links to flowering time pathways (NRT1,

AT59, ANR1, ATX2, and HKT1) as well as genes involved in zinc transport and leaf and stem morphogenesis (ZIP5 and ANAC036). An especially strong candidate is ATX2, an H3K4-specific methyltransferase that affects flowering time in a FRI-dependent manner by regulating levels of FLC mRNA (Pien et al. 2008; Saleh et al. 2008; Shafiq et al. 2014). In the natural population, we identified a missense variant in this gene, L125F, segregating at 66.67% frequency and associated with a 4-day reduction in flowering time in the FRI derived background (Tukey test FRI-ATX2, Der-Der/Der-Anc: diff $= -5.43$ days, P-value $= 4.82 \times 10^{-10}$; beta from genome-wide efficient mixedmodel association (GEMMA) LMM with FRI as a covariate $= -3.8$ days, $P = 0.0089$). Our results agree with previous mutant analyses showing a strong effect of ATX2 on the number of leaves at bolting (as a proxy for flowering time) (Shafiq et al. 2014). Overall, in Santo Antão, we found that candidate loci comprise genes involved in core flowering pathways as well as those with less direct links to flowering time.

Previous studies that used mapping populations derived from a cross between the A. thaliana lines Cvi-0 (derived from a plant from Santo Antão) (Fulgione et al. 2022) and Ler-0 (derived from a Eurasian plant) (Alonso-Blanco, Peeters, et al. 1998; Alcázar et al. 2014) identified a CVI haplotype that resulted in daylength insensitive early flowering (Alonso-Blanco, El-Assal, et al. 1998). This haplotype was initially designated the EDI-Cvi-0 (Early Day-length Insensitive) allele and was later localized to a specific missense variant (V367M) in the Cryptochrome 2 blue light photoreceptor, CRY2 (Alonso-Blanco, Peeters, et al. 1998; El-Din El-Assal et al. 2001; Keurentjes et al. 2007). When introgressed into a European background (Ler-0), EDI-Cvi-0 has a large impact on flowering time, reducing the time to flower such that leaf number at flowering is 18–22 leaves less in short day conditions (Alonso-Blanco, El-Assal, et al. 1998). Subsequently, using transgenic allelic exchange, it was shown that the Cvi-0 specific V367 variant was sufficient to explain the difference in flowering time between the Cvi-0 allele relative to Ler-0 (El-Din El-Assal et al. 2001). This represents one of the earliest and best-validated cases of a natural variant's functional impact on trait variation.

We specifically examined the CRY2 367M allele in the natural population, which we found was private to and segregating at high frequency in Santo Antão (90%). In GWAS, we found neither a significant nor a suggestive effect of this variant in the natural population (LMM in GWAS, P-value $= 0.945$; fig. 3, supplementary figs. S1 and S2, Supplementary Material online), even after correction for the effect of FRI 232X (LMM, P-value $= 0.318$). The discrepancy between our results and previous findings could have been due to a difference in the conditions in our experiment. To test this hypothesis, we grew the near-isogenic lines (NILs) that were previously used to map the EDI-Cvi-0 allele (Alonso-Blanco, El-Assal, et al. 1998; Keurentjes et al. 2007) in our growth conditions. These NILs contain the EDI-Cvi-0 (CRY2) locus introgressed into
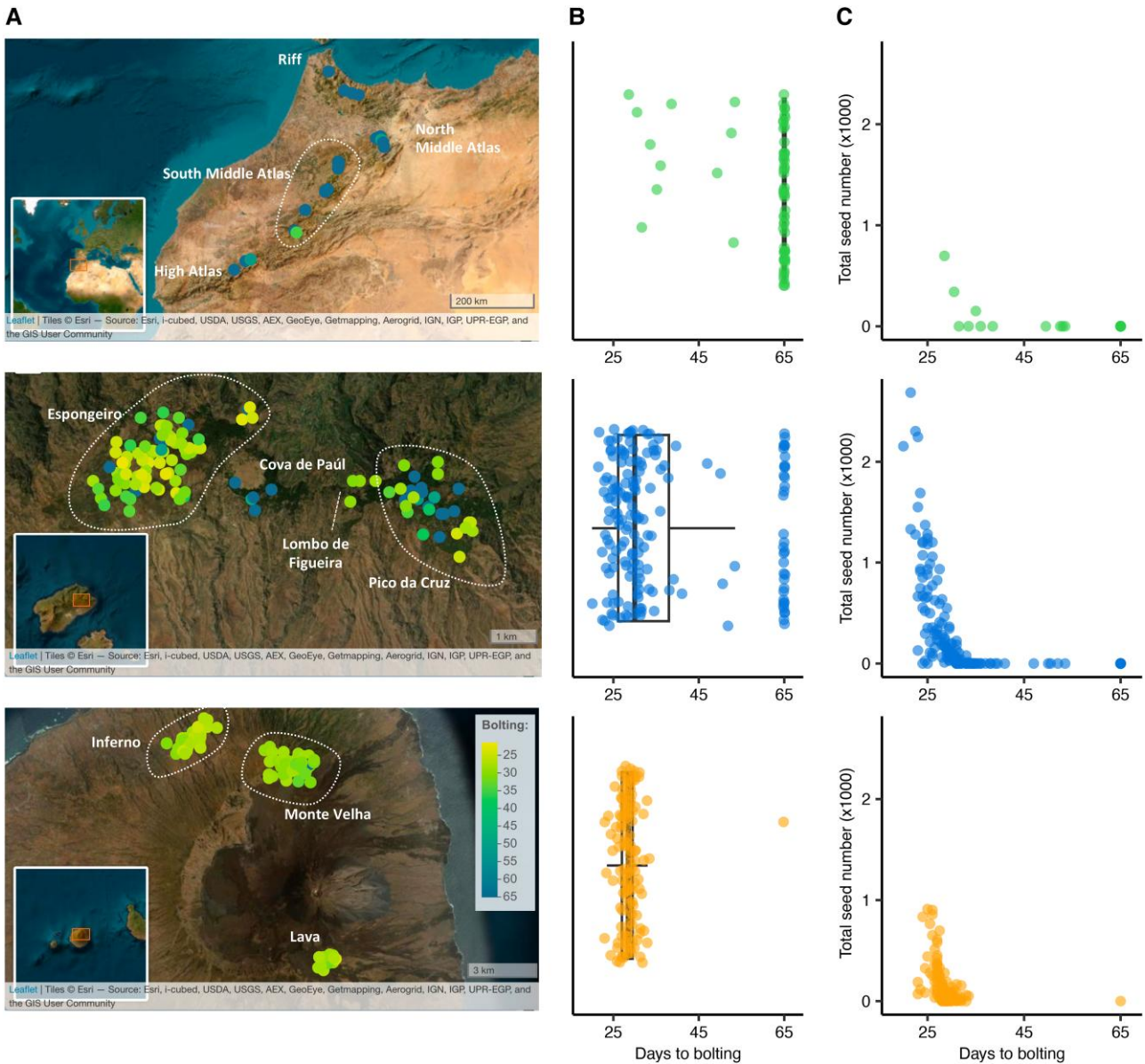
**Fig. 2.** Flowering time varies in the populations and is correlated with fitness. Results are shown from top to bottom for Moroccan, Santo Antão, and Fogo populations, respectively. (*A*) Spatial distribution of variation in flowering time. (*B*) Phenotypic distribution of flowering time per population. Each dot corresponds to the median across four replicates per line. (*C*) Scatter plot showing the time to first flowering versus seed production, a proxy for fitness.

the European (Ler-0) background (El-Din El-Assal et al. 2001; Keurentjes et al. 2007). This experiment recapitulated the previously reported reduction in flowering time due to *EDI-Cvi-0*, showing that in the European genetic background, even under simulated CVI conditions, *CRY2* 367M reduces flowering time (supplementary fig. S2, Supplementary Material online) (Alonso-Blanco, Peeters, et al. 1998; El-Din El-Assal et al. 2001). When adding *CRY2* to a phenotype-prediction model alongside the other associated variants identified here, we found no evidence for a marginal or interaction effect. An epistatic effect due to a variant that is fixed in Santo Antão relative to Ler-0 is still a possibility. Another possibility is that a variant not carried by Cvi-0 but in LD with *CRY2* 367M that

increases flowering time interferes with the detection of *CRY2* V367M in GWAS. To investigate this, we identified variants in high LD with *CRY2* V367M in Santo Antão. Of the ten variants with $r^2 > 0.8$ with *CRY2* V367M, we found no clear flowering time candidate loci represented. Still, we cannot rule out the possibility that variants in LD with *CRY2* V367M in the natural population could affect our power to detect *CRY2* V367M in GWAS. Further work will be needed to understand the role of the *CRY2* variant on trait variation in Santo Antão and to determine with certainty whether the lack of signal there is due to a difference in the genetic background between Cvi-0 and Ler-0.

Next, we conducted GWAS in the Fogo population. Previously, we showed that the reduction in flowering

time in Fogo is largely due to a fixed major effect mutation in the core vernalization pathway (*FLC* 3X) (Fulgione et al. 2022). In this study, GWAS revealed several additional smaller-effect associations with flowering time. We found signals of association near *PAPP2C* (AT1G22280), the gene that encodes a phosphatase that interacts with phytochromes A and B (Phee et al. 2008), and *OXS2* (AT2G41900), which encodes a stress-induced transcription factor that interacts with SOC1, FT (the gene symbol for FLOWERING TIME LOCUS T), and FD (the gene symbol for FLOWERING LOCUS D; three proteins known to induce the transition to flowering) (Blanvillain et al. 2011). The strongest association signal in Fogo overlapped AT4G02480, which encodes an AAA-type ATPase protein. Although its specific function is unknown, AT4G02480 has been shown to interact with the core flowering time

proteins FKF (the gene symbol for FLAVIN-BINDING, KELCH REPEAT, F BOX 1;) and GI (the gene symbol for GIGANTEA), which regulate CO (the gene symbol for CONSTANS) protein stability for photoperiod control of flowering (Song et al. 2014). We also identified associations at two other candidate genes belonging to gene families associated with flowering time in Santo Antão: *ATX1* (AT2G31650) and *NPF5.15* (AT1G22570). ATX1 is a histone-lysine *N*-methyltransferase involved in the formation, placement, and identity of floral organs, and the epigenetic control of *FLC* (Pien et al. 2008). ATX1 is also involved in seed germination, stomatal aperture, water loss, and sensitivity to dehydration stress (Ding et al. 2011, p. 1). In NPF5.15, a protein involved in nitrate and hormone transport (Léran et al. 2014), a nonsense variant is associated with flowering time in Fogo.

In the Moroccan GWAS, a large number of genomic loci were implicated in flowering time (fig. 3). These include nonsynonymous variants in genes directly involved in flowering time as well as loci involved in other traits. Loci directly implicated in flowering time include AT3G19140 (*DNF* L7V), AT2G27250 (*CLV3* M1fs), and AT1G22950 (S131T/G). *DNF* is an E3 ligase that represses *CO* and is crucial to distinguish between long and short days, preventing flowering in short days (Morris et al. 2010; Morris and Jackson 2010). *CLV3* is one of the three *CLAVATA* genes controlling the size of the shoot apical meristem in *Arabidopsis* and it regulates shoot and floral meristem development. AT1G22950 participates in the epigenetic repression of several members of the MADS-box transcription factor family during vegetative development via histone modification. It directly targets *FLC* and mutants disrupt the cold-induced

**Table 1.** Heritability of Flowering Time Under CVI Simulated Conditions and the Inferred Number of Loci Contributing to Trait Variation Using Three Complementary Approaches.

| | Heritability | | Number of loci | | |
|---|---|---|---|---|---|
| | $H^2 = \frac{var_{line}}{var_{line}+var_{residual}}$ [a] | PVE[b] (95%CI)[c] | BSLMM[d] | Clump | LocalScore |
| Morocco | 80.10 | 99.96 (99.8–99.99) | 39 | 6666 | 353 |
| Santo Antão | 85.28 | 98.3 (97.6–98.8) | 10 | 104 | 6 |
| Fogo | 51.10 | 95.2 (92.5–97.6) | 3 | 159 | 26 |

[a]Broad-sense heritability calculated across replicates (in %).
[b]Percentage of phenotypic variance explained by relatedness (narrow-sense heritability, in %).
[c]95% CI across ten runs in BSLMM (in %).
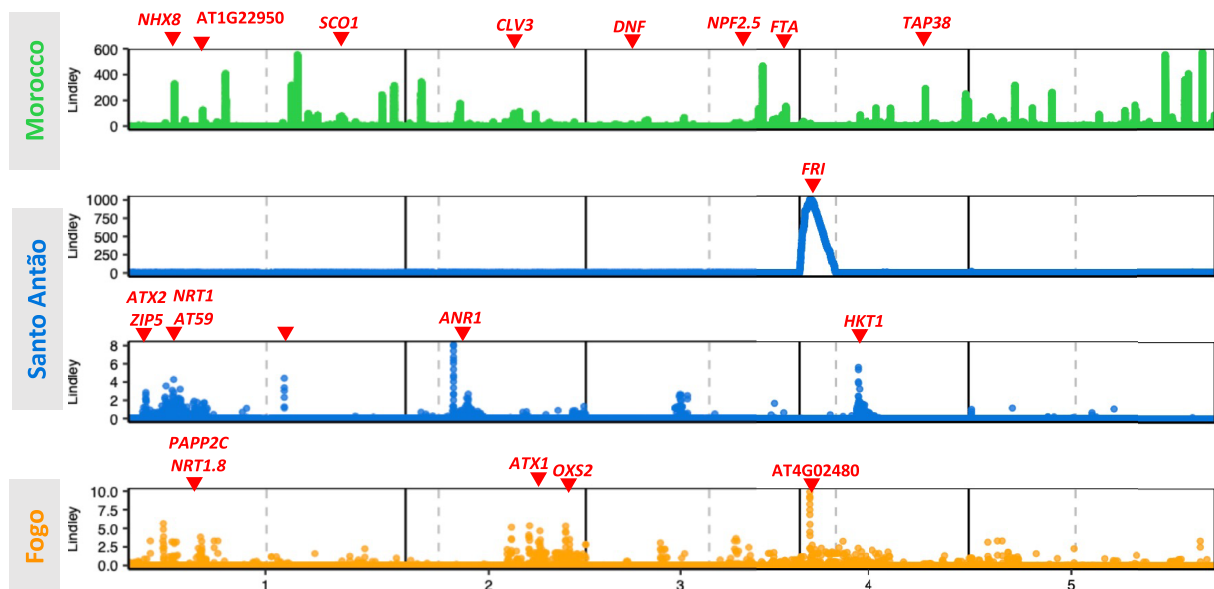[d]Bayesian sparse LMM.



**Fig. 3.** GWAS for flowering time. Manhattan plots for the three populations (Morocco, Santo Antão, and Fogo) show associations with flowering time under CVI simulated conditions. The chromosome position is shown on the x-axis and the Lindley score from the local score approach is shown on the y-axis. Candidate genes are denoted by arrows. For Santo Antão, the top panel shows GWAS results with all genotyped markers and the bottom panel shows GWAS results with *FRI* K232X as a covariate.

Polycomb-mediated silencing underlying vernalization (Bloomer and Dean 2017).

Loci with GWAS signals that are not directly linked to the flowering-time pathway may indirectly contribute to variation in flowering time. In Morocco, we identified signals for variants in genes involved in mineral uptake and processing (*NPF2.2*, *NPF2.5*, and *NHX8*), photosynthesis and growth (*TAP38* and *SCO1*), and response to ultraviolet radiation and drought stress (*MPK1* and *FTA*). Overall, GWAS results from the three populations included a mix of loci involved directly in flowering time and those involved in other linked pathways.

## The Genetic Architecture of Flowering Time is Less Polygenic in the CVI

We next used the results from GWAS to assess whether trait polygenicity differed between the CVI and Morocco. We hypothesized that the flowering time trait would be less polygenic in the younger, less genetically diverse island populations compared to the older Moroccan populations. Since there is no single gold-standard method to infer trait architecture, we tested this hypothesis by applying three complementary approaches to estimate the number of loci contributing to the trait. These included an approach that estimates trait architecture assuming a mixture of large and small to infinitesimal effects (BSLMM) (Zhou et al. 2013), as well as two distinct approaches to reduce redundancy in GWAS results based on local LD. To estimate the genetic architecture from GWAS results by reducing redundancy due to local linkage (LD) in these, we applied 1) a clumping algorithm that maintains only the variants with $r^2 < 0.8$ (Purcell et al. 2007) and 2) the "local score" approach described above (Fariello et al. 2017; Bonhomme et al. 2019).

The estimated numbers of loci contributing to the trait varied across methods, but all approaches resulted in the same pattern: many more loci underlie flowering time variation in the Moroccan population compared to the island populations (fig. 3, table 1). The estimated number of loci impacting the trait tended to be lowest with the BSLMM approach. This may be due to its integration of sparsity and shrinkage across the set of genome-wide associated loci in contrast to the other approaches that only remove local redundancy due to LD. The estimate from BSLMM may therefore be more conservative, but it is less likely to include causative loci than the other approaches. The local score method resulted in an intermediate number of loci compared to the extremes from the other two methods. Below, we use the local score results to investigate the evolutionary history of sets of trait-associated loci.

A limitation of this study was the relatively small sample size for GWAS in the Moroccan population ($n = 62$). Although we likely captured most genetic variation segregating at an appreciable frequency in CVI, we are underpowered in Morocco, where genetic variation is very high. However, the smaller sample size in Morocco should lead to an underestimate of the number of loci contributing to the trait in Morocco, so our finding that the flowering time trait is more polygenic in Morocco compared to the island populations should be robust to sample size differences.

## Patterns in the Cape Verde Populations are Consistent With Strong Directional Selection Under a Fisher–Orr Adaptive Walk Model

Under very strong directional selection, an exponential distribution of effect sizes is expected, with the largest-effect variants arising early in the adaptive walk (Orr 1998, 2005). In contrast, under a model of stabilizing selection or weak directional selection, many loci with a uniform distribution of effect sizes are expected to contribute to trait variation (Orr 1998; Barton et al. 2017; Koch and Sunyaev 2021; Simons et al. 2022). We investigated the distribution of effect sizes (ß estimated in GEMMA with an LMM) in each population using one representative variant per candidate locus (identified with the local score approach and LD-pruned to remove SNPs with $r^2 > 0.5$) (fig. 4, supplementary table S1, Supplementary Material online). We manually added information about *FLC* R3X, which is fixed in Fogo. First, we compared the effect size distributions for the three populations and found that the Moroccan distribution is significantly different from the two islands' [Mann–Whitney (MW) test, Morocco–Santo Antão: *P*-value = 0.033, Morocco–Fogo: *P*-value = $1.3 \times 10^{-6}$]. Then, to investigate whether these effect sizes fit a uniform distribution, as predicted under Fisher's infinitesimal model (Fisher 1918), or an exponential distribution, as predicted by the Fisher–Orr geometric model of adaptation (Fisher 1930; Orr 1998), we compared the fit of these two classes of distributions to the effect size distribution per population using a maximum likelihood approach. In Morocco, the distribution of effect sizes best fit a uniform distribution (Aikake Information Criterion [AIC]: exponential: 1810.13, uniform: 1568.97), while in Santo Antão and Fogo, effect size distributions best fit an exponential distribution (AIC $_{Santo\ Antão}$: exponential: 30.21, uniform: 31.58; AIC $_{Fogo}$: exponential: 127.94, uniform: 128.96). Overall, our results are consistent with evolution dominated by stabilizing selection in the Moroccan population and with strong directional selection in the CVI populations.

We next investigated the relationship between allele frequencies and their associated effects to determine whether derived alleles tended to increase or decrease flowering time. In Morocco, we found that effect size was positively correlated with derived allele frequency (Pearson's $R = 0.34$, *P*-value = $1.98 \times 10^{-8}$; fig. 5). When we specifically tested the effect for loci that increased versus decreased flowering time, we found that an increase in flowering time (i.e., later flowering) was significantly positively correlated with derived allele frequency (Pearson's $R = 0.65$, *P*-value $< 2.2 \times 10^{-16}$). On the other hand, for loci that decreased flowering time, we observed a negative correlation between flowering time and derived allele frequency (Pearson's $R = -0.72$, *P*-value = $1.44 \times 10^{-15}$). On the islands, we found that alleles that accelerate flowering are mainly present at high frequencies while alleles that
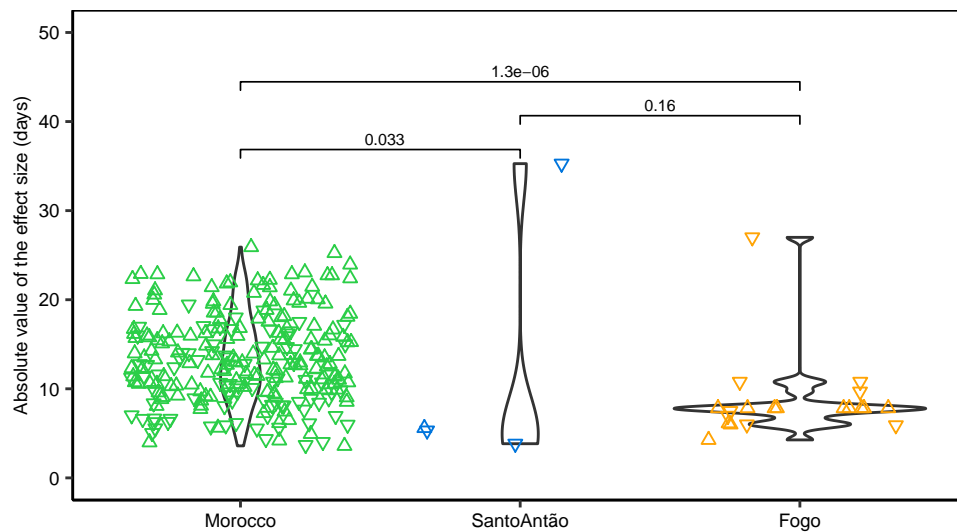
Fig. 4. Effect size distributions of flowering time-associated loci. The absolute value of the effect size in days (y-axis) per population (x-axis). Each dot represents one SNP per candidate locus identified with the local score approach and with a $r^2 < 0.5$. Triangles pointed up represent variants with an estimated positive effect size, delaying flowering, while triangles pointed down represent variants with an estimated negative effect size, reducing flowering time. Although *FLC* R3X is fixed in Fogo, and not identified through GWAS, it is included here for completeness. *P*-values shown are from MW tests.

delay flowering are present at low frequencies, suggesting a history of directional selection. With so few variants, statistical analysis of correlations had low power, especially in Santo Antão, but the magnitudes of associations were strong in both cases (Santo Antão: Pearson's $R = -0.74$, *P*-value = 0.263, Fogo: Pearson's $R = -0.76$, *P*-value = $9.84 \times 10^{-5}$; fig. 5). Our results are consistent with a predominant force of stabilizing selection acting to maintain late flowering in Morocco and parallel directional selection for earlier flowering in the islands. Consistent with results from fitness experiments (Fulgione et al. 2022), these results imply that in *Arabidopsis* populations that colonized Cape Verde, selection favored alleles that reduced flowering time because it allowed the populations to reproduce even with the shorter growing season. Conversely, in Morocco, where the growing season is longer, alleles that cause flowering to be later would tend to be beneficial.

The near-complete loss of variation that occurred with the colonization of the two CVI provides a rare opportunity to examine how variation in a quantitative trait builds up over time. To this end, we next estimated the ages of loci identified in GWAS together with *FRI* 232X and *FLC* 3X using a coalescent approach (Speidel et al. 2019). On both islands, older flowering time-associated variants are at higher frequencies than younger variants (Santo Antão: Pearson's $R = 0.999$, *P*-value = 0.000461; Fogo: Pearson's $R = 0.659$, *P*-value = 0.0008517; fig. 6, supplementary table S2, Supplementary Material online). In Santo Antão, the oldest candidate variant, with an age estimated at approx. 2,700 years, causes a premature truncation of *FRI* K232X and was soon followed by a moderate effect missense mutation in *ATX2* (L125F, 2,300 years ago). Smaller effect loci, mainly modifiers, appeared more recently (in the last 200 years). In Fogo, *FLC* R3X arose approx. 3,000 years ago, followed by variants with smaller effects and moderate predicted functional effects in the last 1,000 years, with most arising in the last 500 years.

Overall, these results are consistent with a model where flowering time decreased on the archipelago through an adaptive walk, with large-effect size mutations arising first and quickly increasing in frequency (*FRI* K232X and *FLC* R3X), followed more recently by smaller effect size alleles. The most recently arising and lowest frequency loci often act to increase the time to flowering. These could be fine-tuning flowering time as expected under a Fisher–Orr model or they could be deleterious variants that have not yet been purged from the population.

## Discussion

Understanding trait architectures and the factors that shape them is a central goal in evolutionary biology, and it has applications in medical genetics, conservation genetics, and breeding (Falconer and Mackay 1996; Walsh and Lynch 2018; Bomblies and Peichel 2022; Charlesworth et al. 2022). Consistent with the infinitesimal model (Fisher 1918), highly polygenic architectures have been found in cases where population diversity is high and stabilizing selection or weak directional selection is the dominant evolutionary force acting on the population (Rockman 2012). For example, in *Drosophila*, body size variation has been attributed to hundreds to thousands of genomic regions (Turner et al. 2011; Pallares et al. 2023), and similarly, in mice, thousands of loci are needed to explain variation in body size (Reed et al. 2008). In a European human population, a study of more than 5 million individuals determined 12,111 associated SNPs could explain nearly all SNP heritability in height (Yengo et al. 2022). Expression QTL studies similarly tend to find that overall expression is highly polygenic (Josephs et al. 2015; He et al. 2016; Liu et al. 2019; Võsa et al. 2021). Overall, there is considerable evidence that traits in natural populations are often polygenic.

However, when selection is very strong and the population is mutation-limited with respect to the selected trait, large-effect variants may be important in adaptation, causing the infinitesimal model to break down (see Barton et al. 2017, p. 57). Extending on Fisher's geometric model, Orr
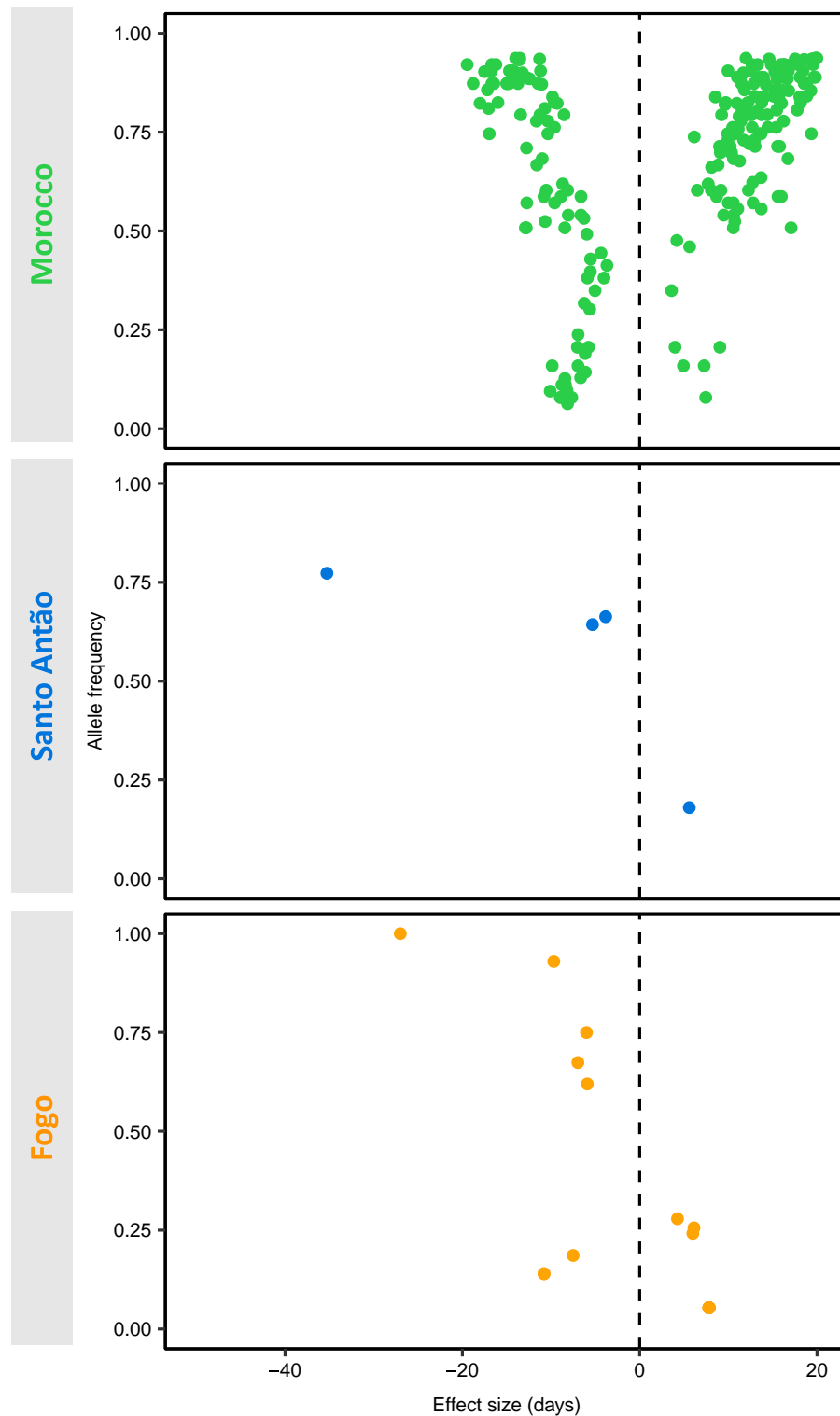
**FIG. 5.** Relationship between allele frequency and effect size for flowering time-associated variants in the three populations. Effect size in days (x-axis) for *FRI* 232X, *FLC* 3X, and each SNP tagging a candidate locus (each dot) and its respective allele frequency on the population (y-axis). A negative effect size corresponds to early flowering and a positive effect size to late flowering.

produced a model of adaptation in which a mutation-limited population facing a sudden environmental shift follows an adaptive walk toward the new fitness optimum (Orr 1998, 2002). Empirical support for the Fisher–Orr geometric model comes from a broad range of traits and species (Dittmar et al. 2016; Connallon and Hodgins 2021). These include cases in microbial evolution (Arjan et al. 1999; Barrick et al. 2009; Schoustra et al. 2009; Szendro et al. 2013; Good et al. 2017); "industrial melanism" in *Biston betularia*, which is caused by a single large-effect locus that results in darker pigmentation and which rose to high frequency in highly industrial areas (Saccheri et al. 2008; van't Hof et al. 2011, 2016); color patterning in mice living on pale coastal sand dunes versus dark lava or mainland environments, which is determined by variation in two genes, *Mc1r* and *Agouti* that together explain
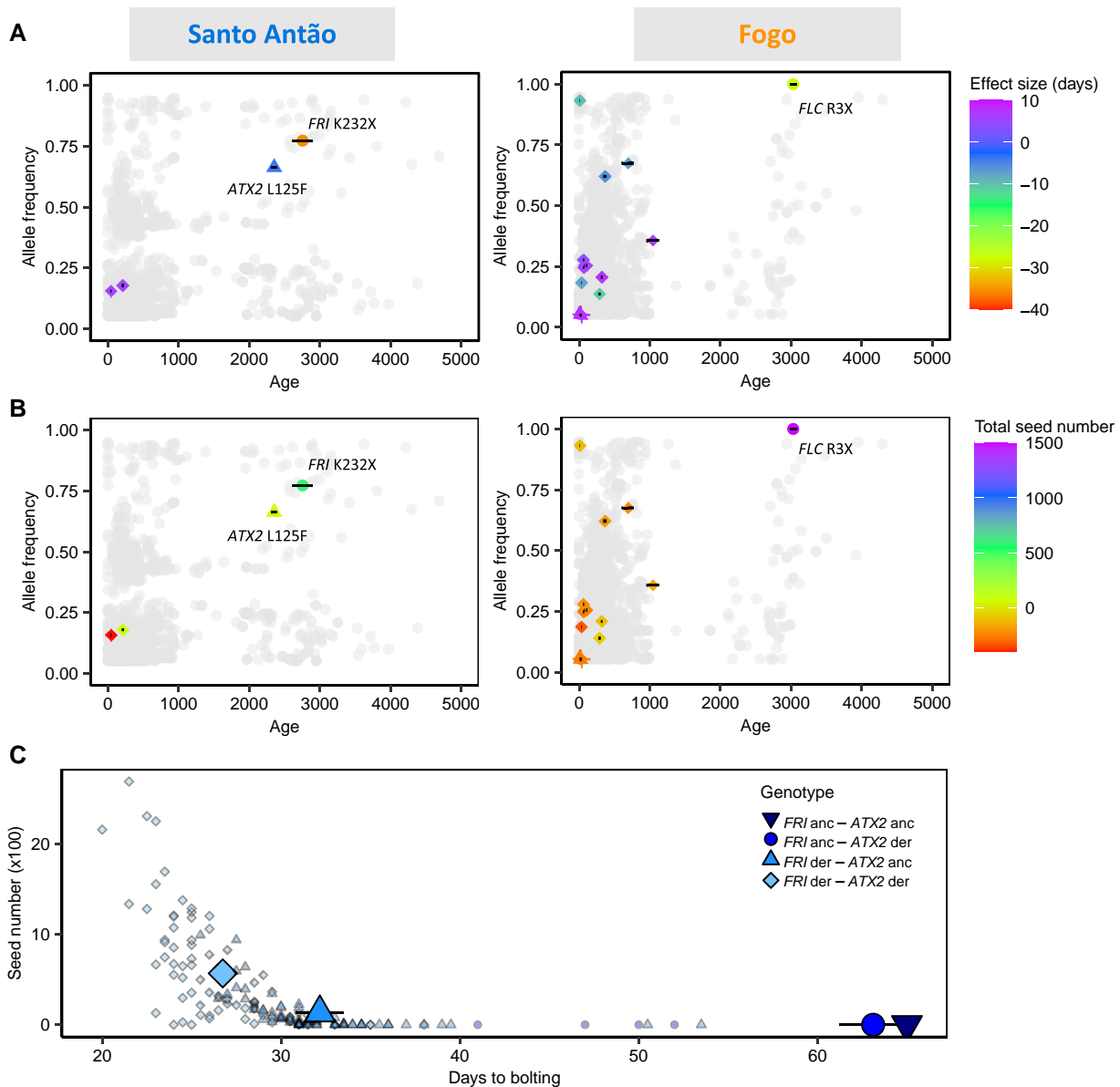
Fig. 6. Relationship between age and allele frequency of loci implicated in flowering time and fitness in CVI. In (A and B), age estimates (in years) versus allele frequency for associated loci are shown in comparison to the genomic background. Colored SNPs represent associated variants, with colors matching their estimated effect sizes in days and seed number (A and B, respectively), and shape their predicted impacts (based on SnpEff annotation; circles are high impact, triangles moderate, diamonds are modifiers, and cross low effect variants). Each gray dot represents one SNP in an LD-pruned genome. The left panels refer to Santo Antão and the right panels to Fogo. Horizontal black lines on the associated variants represent 95% CI of estimated age. (C) Effects of allelic combinations between the two major candidates from Santo Antão, FRI, and ATX2, respectively, on time to flower (x-axis) and fitness (y-axis). Each small dot represents one line from the Santo Antão natural population, and large symbols the average per genotype category.

~40% of variation in pigmentation (Nachman et al. 2003; Hoekstra et al. 2006; Steiner et al. 2007; Weber et al. 2013); body shape variation between marine and freshwater three-spine stickleback fish, which is controlled by two large-effect loci and a few more small effect loci (Shapiro et al. 2004; Colosimo et al. 2005; Chan et al. 2010; Jones et al. 2012; Rogers et al. 2012; Peichel and Marques 2017; Schluter et al. 2021); and beak size variation in Darwin's finches reflecting dietary specialization, which is determined by up to six loci of large effect that explain up to 46% of the phenotypic variation (Lamichhaney et al.

2015; Enbody et al. 2022). In each of these examples, a population adapted to a distant optimum, using variants that had appreciable effects and explained a large proportion of the total phenotypic variance, similar to the island populations we studied here.

It is often implicitly assumed that trait architectures will be similar across populations. However, within a species, the demographic and selective histories of individual populations may vary, which can impact the genetic architectures of traits (Gomulkiewicz and Houle 2009). In this study, we asked whether the genetic architecture of a trait

that contributes to adaptation in island-colonizing populations differs from that in the closest continental outgroup population. We compared the genetic architecture of flowering time in *Arabidopsis thaliana* populations that recently (ca. 4–5 kya) colonized a novel environment defined by a short growing season with their closest mainland outgroup. We hypothesized that strong directional selection in the Cape Verde populations combined with a relatively recent founder event would result in a pattern that is consistent with Orr's model of an adaptive walk (Orr 1998). Conversely, in the much older continental Moroccan population, which coalesces at approx. 1 mya, we hypothesized that the fitness optimum would largely have been reached and that patterns would be consistent with stabilizing selection or weak directional selection, as expected under an infinitesimal model of evolution (Fisher 1918; Barton et al. 2017).

We found several lines of evidence that the CVI populations were evolving under strong directional selection consistent with a Fisher–Orr geometric model. First, we found that polygenicity of the genetic architecture of the flowering time was reduced in the Cape Verde populations relative to the Moroccan population. Second, we found that effect sizes of trait-associated variants in the CVI populations were exponentially distributed, in contrast to the uniform distribution of effect sizes in Morocco. Although an exponential distribution of effects may be possible under neutral evolution (Robertson 1967; Cotto and Day 2023), other aspects of our results provide further evidence for strong directional selection. Third, based on reconstructed ages of the trait-associated variants, those implicated in large flowering time reductions arose first in CVI (*FLC* 3X, *FRI* 232X, and *ATX2* 125F) and smaller-effect variants arose more recently. Finally, the strong effect of reduced flowering time and the specific large-effect variants on fitness further support the role of these variants in adaptation. Taken together, the combination of an exponential distribution of effect sizes and strong fitness effects of *FLC* 3X in Fogo, and *FRI* 232X and *ATX2* 125F in Santo Antão provide evidence for an adaptive walk, consistent with a Fisher–Orr model.

Although we found evidence for an adaptive walk in the CVI populations, there are at least two ways in which the populations do not perfectly fit the Fisher–Orr model. First, *FRI* 232X alone and in combination with *ATX2* 125F has strong effects on flowering time and fitness (fig. 6), but neither variant is fixed in Santo Antão (supplementary fig. S3, Supplementary Material online). This could potentially be explained partly by population structure within the island, which likely developed early (Fulgione et al. 2022), and partially in response to a spatially variable climate. Specifically, the humid trade winds that provide much of the total moisture to the islands produce an east–west gradient in humidity during the wet season (Brochmann et al. 1997; Fulgione et al. 2022; Elfarargi et al. 2023). We are currently extremely limited in available models of quantitative traits, with existing models tending to assume random mating. Models that

integrate population structure *and* that are realistic, that are informed by real data, will be important to advance our ability to make nuanced comparisons between models and data. Second, while there is strong evidence based on fitness in simulated CVI conditions (fig. 6) that the early arising variants were adaptive, the adaptive relevance of the later-arising flowering-time-associated variants is less clear. The youngest and lowest frequency variants are often associated with delayed flowering time (fig. 6). Among these, there were no clear candidate loci belonging to pathways directly connected to flowering time. Instead, we identified candidates involved in more peripheral pathways such as nutrient/metal uptake and regulation (e.g., *NRT1*, *NRT1.8*, *ANR1*, *ZIP5*), and light sensing (e.g., *PAPP2C*, *AT4G02480*). As these other loci were expected to affect flowering time only indirectly and thus only slightly adjust flowering time (Mouradov et al. 2002; Andrés and Coupland 2012; Sanagi et al. 2021), they may have a role in adaptation to other selective pressures such as the volcanic soil, the near-constant photoperiod, or the higher light intensity. Alternatively, these loci may simply represent segregating weakly deleterious mutations. Our results point to the need for studies that examine patterns generated under more diverse and realistic evolutionary models.

The omnigenic model builds on Fisher's infinitesimal model, proposing that highly polygenic complex trait variation is shaped largely by genes in peripheral rather than in core trait pathways (Boyle et al. 2017). According to this model, the high degree of connectivity in biological networks leads to the situation where most expressed genes are only a few steps from the nearest core gene and thus affect a trait through their network interactions (Boyle et al. 2017). Under this model, association signals from pathways other than the flowering-time pathway would be expected due to the complex network of molecular interactions between the flowering-time pathway and more peripheral pathways. Our results are consistent with the idea that populations far from the adaptive optimum are likely to initially move closer using large-effect variants in core genes. However, as the population moves nearer to the adaptive optimum or reaches it, more peripheral genes could predominate. This would be consistent with the idea that traits evolving at or near the optimum (i.e., evolving under stabilizing selection) may often fit an omnigenic model (Boyle et al. 2017), with the recently arising small effect variants largely representing slightly deleterious mutations that have not yet been removed by selection.

We did not find evidence of association for the *CRY2* V367M variant (in the *EDI-Cvi-0* allele) carried by Cvi-0 that was previously shown to have an effect on flowering time (Alonso-Blanco, El-Assal, et al. 1998; El-Din El-Assal et al. 2001, 2003). Common alleles at this locus have also been connected to flowering time variation in Eurasia (Olsen et al. 2004). The main function of *CRY2* is as a light sensor and variation in this gene have been shown to affect many traits in addition to flowering time, including light and temperature responsiveness (Balasubramanian et al.

2006; Sanchez-Bermejo et al. 2015), chromatin condensation (Tessadori et al. 2007), fruit length, ovule number per fruit, percentage of unfertilized ovules and silique number (Botto et al. 2003; el-Assal et al. 2004; Fournier-Level et al. 2013). Although we recapitulated the previous finding of reduced flowering time in the NILs that carried EDI-Cvi-0 in a Ler-0 background under our experimental conditions, we detected no significant effect of CRY2 V367M in the CVI natural population. This discrepancy could be due to an epistatic effect that abrogates the CRY2 effect on flowering time in the CVI genetic background. As we did not find evidence for an epistatic interaction with other segregating variations in Santo Antão that we identified in GWAS, the most likely scenario may be that genetic interaction with a variant fixed in the ancestor of CVI relative to Ler-0 is responsible for the lack of signal in the natural population. Regardless, the CRY2 V367M allele is at high frequency (90%) in the natural population of Santo Antão, suggesting that it may have been positively selected. Upon colonization from North Africa, CVI Arabidopsis would have experienced a sudden shift in several environmental factors and the CRY2 V367M variant may have been selected due to its effect on other traits. Further work on the effects of CRY2 V367M on other traits is needed to fully understand its function and relevance under natural conditions. More generally, this finding demonstrates the importance of exercising caution when extrapolating effects observed in one genetic background to other backgrounds. It further points to the need to select appropriate genetic backgrounds, informed by evolutionary history, when examining gene and allelic functions.

Overall, this study provides an empirical example of how genetic architecture built up over time in two relatively young populations that experienced strong directional selection. In addition, it provides insights into how trait architecture is shaped by the combination of demographic and selective forces in natural populations. These results will serve as a useful case for comparison in future work that investigates genetic architectures across traits, populations, and species.

## Materials and Methods

### Genomic Data and Population Structure Analyses

All genomic data on the Moroccan and CVI natural populations were previously published (European Nucleotide Archive codes PRJEB24044 and PRJEB39079, respectively) (Durvasula et al. 2017; Fulgione et al. 2022). We used SNP and InDel variants called with the genomic analysis toolkit (GATK) pipeline and published in (Fulgione et al. 2022). The neighbor-joining tree was produced using the R package ape and principal component analysis was conducted using the flag <−pca> in PLINK v.1.90 (Purcell et al. 2007).

### Experimental Conditions and Germplasm

We propagated plants from two CVI, Santo Antão ($n = 174$) and Fogo ($n = 129$), together with plants from

Morocco ($n = 62$), in a custom Bronson growth chamber (Brennan et al. 2014; Fulgione et al. 2022). Before sowing, seeds were stratified in the dark in Petri dishes on water-soaked filter paper for one week at 4 °C, and then sown in $7 \times 7$ cm pots, containing a standard potting compost mix, supplemented with iron. Four replicates of each accession were sown and organized in a randomized block design. Since flowering time heritability tends to be high, and space within the growth chamber was limited, we opted to include a limited number of replicates and instead maximize the number of accessions. Further, compared to a study design with higher line-replication, this design provides high replication at the level of individual alleles (variants) across the population. The chamber was set to track hourly environmental data (temperature, humidity, photoperiod, and precipitation) from Cape Verde, simulating the growing season. Following the precipitation loggers in the field, water was withheld for 26 days after sowing (Fulgione et al. 2022).

Bolting time, that is the number of days from sowing until the appearance of the differentiated floral bud, was scored per individual and the median across replicates was taken as the phenotype per genotype. The term "flowering time" is used throughout the paper to represent the scored bolting time, as the latter is a proxy for the former. At the end of the experiment, the total number of produced seeds was assessed, as a proxy for fitness.

We also grew two Cvi-0×Ler-0 NILs—LCN1–2.5 and LCN1-2.8—and the two parental lines Cvi-0 and Ler-0 to assess the effect of CRY2—an a priori candidate–, under CVI simulated conditions. Each line was propagated in four replicates in a randomized block design. Phenotyping was conducted as indicated above.

### Genome-Wide Association Mapping and Genetic Architecture Inference

Broad-sense heritability ($H^2$) was calculated in R using a linear mixed-effect model that accounts for the block design, using the function lmer() from the package lme4 (Bates et al. 2015) and narrow-sense heritability (chip heritability) was calculated in GEMMA using BSLMM (Zhou et al. 2013) (see below). To map flowering time, we conducted GWAS using GEMMA v.0.94 (Zhou and Stephens 2012), with some a posteriori modifications (see below). We used SNP and InDel variants called with the GATK pipeline and published in (Fulgione et al. 2022). All input files were generated using VCFtools v.0.1.14 (Danecek et al. 2011) and PLINK v.1.90 (Purcell et al. 2007), and the median bolting time across replicates per genotype was used as the phenotype.

To map the genetic basis of flowering time variation, we used a univariate LMM that accounts for population structure with a centered kinship matrix <-gk 1> in GEMMA v.0.94 (Zhou and Stephens 2012). We used the flag <-lmm 4>, which estimates beta per marker as the effect size of an allele, and used the likelihood ratio test to assess evidence for association. For mapping in Santo Antão with

*FRI* K232X as a covariate, the flag <-c> was used with a genotype file coded with 0 and 1 for the two alleles. The same approach was used to obtain effect size estimates per marker for seed number.

To infer trait genetic architecture, we used three methods: BSLMM, Lindley (local) scores, and clumping. First, BSLMM (Zhou et al. 2013) was run in GEMMA (Zhou and Stephens 2012). Markov chain Monte Carlo (MCMC) was run with 10,000,000 sampling steps and 2,500,000 burn-in iterations. Median and 95% confidence interval (CI) for PVE and the number of variants with sparse effects (n.gamma) were calculated across ten runs. Second, to compute Lindley scores, also referred to as "local scores" (Fariello et al. 2017; Bonhomme et al. 2019), from the LMM output, we used available scripts (from https://forge-dga.jouy.inra.fr/projects/local-score/documents) (Bonhomme et al. 2019). Each significant zone was considered a separate candidate locus. Third, clumping of LMM results based on LD was performed using PLINK v.1.90 (Purcell et al. 2007) with the flags <–clump>, <–clump-p1 0.01>, <–clump-kb 1000>, and <–clump-r2 0.8>. Clumps were formed around central "index variants' which are significantly associated with flowering time at $\alpha = 0.01$, and by variants that are within 1 Mbp distance and with $r^2 > 0.8$ with the index marker. A phenotype prediction model was fitted to the local score-identified variants, in addition to *CRY2* V367M, using the *aov()* function in R.

### Testing for an Adaptive Walk

For all analyses in this section, only SNP markers were used. These were filtered from the initial variant call format (VCF) file using the command <–remove-indels> in VCFtools v.0.1.14 (Danecek et al. 2011). SNPs adjacent to apparent complex structural variants were eliminated from the dataset, as these variant calls were unreliable. As noted in the Results section, since *FLC* 3X is fixed in Fogo, where it is included in analyses, it was added manually.

For each candidate locus identified with the local score approach, we took one representative SNP and recalculated its allele frequency and effect size based on the GEMMA estimates. For allele frequency calculations, Moroccan alleles were polarized relative to the *A. lyrata* outgroup. Within CVI, alleles were polarized as described previously (Fulgione et al. 2022). Estimates were recalculated whenever necessary. Candidate loci, represented by a single SNP, were then annotated using SnpEff (Cingolani et al. 2012). The set of candidate loci was further pruned to eliminate SNPs in LD, with an $r^2 > 0.5$.

For each CVI candidate locus and for an LD-pruned genome-wide set of loci, we also estimated allele ages. We pruned the genomes in Santo Antão and Fogo using PLINK v.0.19 (Purcell et al. 2007) and the flag <–indep-pairwise 50 10 0.3> to produce a set of representative loci across the genome. Specifically, we iterated through 50-SNP windows 50, shifting the window in 10-SNP steps, and selected from variants in the window with $r^2 > 0.3$.

Then, we estimated allele ages for all candidate SNPs as well as SNPs in the LD-pruned sets in each CVI population using RELATE (Speidel et al. 2019). Ages and 95% confidence intervals were calculated over 200 MCMC runs using Relate´s *TreeViewSample* script. An important note is that RELATE does not allow missing data, so loci with missing data were removed from the analysis. The mutation rate was corrected for missing data across the entire genome and the recombination map was taken from (Salomé et al. 2011) with a correction to assume an outcrossing rate of 5% in the natural populations, based on previous estimates (Bomblies et al. 2010).

## Supplementary material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Data Availability

All scripts are available at https://github.com/HancockLab/CVI-flowering-time-architecture.

## References

Ågren J, Oakley CG, Lundemo S, Schemske DW. 2017. Adaptive divergence in flowering time among natural populations of Arabidopsis thaliana: estimates of selection and QTL mapping. *Evolution.* **71**(3):550–564.

Ågren J, Schemske DW. 2012. Reciprocal transplants demonstrate strong adaptive differentiation of the model organism Arabidopsis thaliana in its native range. *New Phytol.* **194**(4):1112–1122.

Alcázar R, von Reth M, Bautor J, Chae E, Weigel D, Koornneef M, & Parker JE. 2014. Analysis of a plant Complex resistance gene locus underlying immune-related hybrid incompatibility and its occurrence in nature. *PLoS Genet.* **10**(12):e1004848.

Alonso-Blanco C, El-Assal SE-D, Coupland G, Koornneef M. 1998. Analysis of natural allelic variation at flowering time loci in the landsberg erecta and Cape Verde islands ecotypes of Arabidopsis thaliana. *Genetics.* **149**(2):749.

Alonso-Blanco C, Peeters AJM, Koornneef M, Lister C, Dean C, van den Bosch N, Pot J, Kuiper MTR. 1998. Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a Ler/Cvi recombinant inbred line population: AFLP based linkage map of *Arabidopsis*. *Plant J.* **14**(2):259–271.

Anderson JT, Lee C-R, Mitchell-Olds T. 2011. Life history QTLs and natural selection on flowering time in Boechera stricta, a perennial relative of Arabidopsis. *Evolution.* **65**(3):771–787.

Andrés F, Coupland G. 2012. The genetic basis of flowering responses to seasonal cues. *Nat Rev Genet.* **13**(9):627–639.

Arjan G, Visser J, de M, Zeyl CW, Gerrish PJ, Blanchard JL, Lenski RE. 1999. Diminishing returns from mutation supply rate in asexual populations. *Science.* **283**(5400):404–406.

Balasubramanian S, Sureshkumar S, Lempe J, Weigel D. 2006. Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. *PLoS Genet.* **2**(7):e106.

Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, Schneider D, Lenski RE, Kim JF. 2009. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature.* **461**(7268):1243–1247.

Barton NH, Etheridge AM, Véber A. 2017. The infinitesimal model: definition, derivation, and implications. *Theor Popul Biol.* **118**: 50–73.

Barton NH, Keightley PD. 2002. Understanding quantitative genetic variation. *Nat Rev Genet.* **3**(1):11–21.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *J Stat Sof.* **67**:1–48.

Blanvillain R, Wei S, Wei P, Kim JH, Ow DW. 2011. Stress tolerance to stress escape in plants: role of the OXS2 zinc-finger transcription factor family. *EMBO J.* **30**(18):3812–3822.

Bloomer RH, Dean C. 2017. Fine-tuning timing: natural variation informs the mechanistic basis of the switch to flowering in *Arabidopsis thaliana*. *J Exp Bot.* **68**(20):5439–5452.

Bomblies K, Peichel CL. 2022. Genetics of adaptation. *Proc Natl Acad Sci U S A.* **119**(30):e2122152119.

Bomblies K, Yant L, Laitinen RA, Kim S-T, Hollister JD, Warthmann N, Fitz J, Weigel D. 2010. Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of Arabidopsis thaliana. *PLoS Genet.* **6**(3):e1000890.

Bonhomme M, Fariello MI, Navier H, Hajri A, Badis Y, Miteul H, Samac DA, Dumas B, Baranger A, Jacquet C, et al. 2019. A local score approach improves GWAS resolution and detects minor QTL: application to Medicago truncatula quantitative disease resistance to multiple Aphanomyces euteiches isolates. *Heredity (Edinb).* **123**(4):517–531.

Botto JF, Alonso-Blanco C, Garzarón I, Sánchez RA, Casal JJ. 2003. The Cape Verde islands allele of cryptochrome 2 enhances Cotyledon unfolding in the absence of blue light in Arabidopsis. *Plant Physiol.* **133**(4):1547–1556.

Bouché F, Lobet G, Tocquin P, Périlleux C. 2016. FLOR-ID: an interactive database of flowering-time gene networks in Arabidopsis thaliana. *Nucleic Acids Res.* **44**(Database issue): D1167–D1171.

Boyle EA, Li YI, Pritchard JK. 2017. An expanded view of complex traits: from polygenic to omnigenic. *Cell.* **169**(7):1177–1186.

Brennan AC, Méndez-Vigo B, Haddioui A, Martínez-Zapater JM, Picó FX, Alonso-Blanco C. 2014. The genetic structure of *Arabidopsis thaliana* in the south-western Mediterranean range reveals a shared history between North Africa and Southern Europe. *BMC Plant Biol.* **14**(1):17.

Brochmann C, Rustan ØH, Lobin W, Kilian N. 1997. The endemic vascular plants of the Cape Verde Islands, W Africa. Botanical Garden and Museum, University of Oslo.

Caicedo AL, Stinchcombe JR, Olsen KM, Schmitt J, Purugganan MD. 2004. Epistatic interaction between Arabidopsis FRI and FLC flowering time genes generates a latitudinal cline in a life history trait. *Proc Natl Acad Sci U S A.* **101**(44):15670–15675.

Chan YF, Marks ME, Jones FC, Villarreal G, Shapiro MD, Brady SD, Southwick AM, Absher DM, Grimwood J, Schmutz J, et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a pitx1 enhancer. *Science.* **327**(5963): 302–305.

Charlesworth B, Goddard ME, Meyer K, Visscher PM, Weir BS, Wray NR. 2022. From Mendel to quantitative genetics in the genome era: the scientific legacy of W. G. Hill. *Nat Genet.* **54**(7):934–939.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin).* **6**(2):80–92.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science.* **307**(5717):1928–1933.

Connallon T, Hodgins KA. 2021. Allen Orr and the genetics of adaptation. *Evolution.* **75**(11):2624–2640.

Cotto O, Day T. 2023. A null model for the distribution of fitness effects of mutations. *Proc Natl Acad Sci U S A.* **120**(23):e2218200120.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics.* **27**(15):2156–2158.

Darwin C. 1859. *The origin of species by means of natural selection.* London: J. Murray.

Ding Y, Avramova Z, Fromm M. 2011. The Arabidopsis trithorax-like factor ATX1 functions in dehydration stress responses via ABA-dependent and ABA-independent pathways. *Plant J.* **66**(5):735–744.

Dittmar EL, Oakley CG, Ågren J, Schemske DW. 2014. Flowering time QTL in natural populations of Arabidopsis thaliana and implications for their adaptive value. *Mol Ecol.* **23**(17):4291–4303.

Dittmar EL, Oakley CG, Conner JK, Gould BA, Schemske DW. 2016. Factors influencing the effect size distribution of adaptive substitutions. *Proc R Soc B Biol Sci.* **283**(1828):20153065.

Ducousso A, Guyon JP, Krémer A. 1996. Latitudinal and altitudinal variation of bud burst in western populations of sessile oak (Quercus petraea (Matt) Liebl). *Ann Des Sci Forest.* **53**(2–3): 775–782.

Durvasula A, Fulgione A, Gutaker RM, Alacakaptan SI, Flood PJ, Neto C, Tsuchimatsu T, Burbano HA, Picó FX, Alonso-Blanco C, et al. 2017. African Genomes illuminate the early history and transition to selfing in Arabidopsis thaliana. *Proc Natl Acad Sci U S A.* **114**(20):5213.

el-Assal SED, Alonso-Blanco C, Hanhart CJ, Koornneef M. 2004. Pleiotropic effects of the Arabidopsis cryptochrome 2 allelic variation underlie fruit trait-related QTL. *Plant Biol (Stuttg).* **6**(4): 370–374.

El-Din El-Assal S, Alonso-Blanco C, Peeters AJM, Raz V, & Koornneef M. 2001. A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. *Nat Genet.* **29**(4):435–440.

El-Din El-Assal S, Alonso-Blanco C, Peeters AJM, Wagemaker C, Weller JL, & Koornneef M. 2003. The role of cryptochrome 2 in flowering in Arabidopsis. *Plant Physiol.* **133**(4):1504–1516.

Elfarargi AF, Gilbault E, Döring N, Neto C, Fulgione A, Weber APM, Loudet O, Hancock AM. 2023. Genomic basis of adaptation to a novel precipitation regime. *Mol Biol Evol.* **40**(3):msad031.

Enbody ED, Sendell-Price AT, Sprehn CG, Rubin C-J, Visscher PM, Grant BR, Grant PR, Andersson L. 2022. Large effect loci have a prominent role in Darwin's finch evolution (p. 2022.10.29.514326). bioRxiv. https://doi.org/10.1101/2022.10.29.514326

Exposito-Alonso M, Becker C, Schuenemann VJ, Reiter E, Setzer C, Slovak R, Brachi B, Hagmann J, Grimm DG, Chen J, et al. 2018. The rate and potential relevance of new mutations in a colonizing plant lineage. *PLoS Genet.* **14**(2):e1007155.

Fabian D, Flatt T. 2012. Life history evolution. *Nat Educ Knowl.* **3**(10):24.

Falconer DS, Mackay TFC. 1996. *Introduction to quantitative genetics.* Harlow: Addison Wesley Longman.

Fariello MI, Boitard S, Mercier S, Robelin D, Faraut T, Arnould C, Recoquillay J, Bouchez O, Salin G, Dehais P, et al. 2017. Accounting for linkage disequilibrium in genome scans for

selection without individual genotypes: the local score approach. *Mol Ecol.* **26**(14):3700–3714.

Fisher RA. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Earth Environ Sci Trans R Soc Edinburgh.* **52**(2):399–433.

Fisher RA. 1930. *The genetical theory of natural selection.* Oxford (UK): Clarendon Press. p. 1–302.

Fournier-Level A, Wilczek AM, Cooper MD, Roe JL, Anderson J, Eaton D, Moyers BT, Petipas RH, Schaeffer RN, Pieper B, *et al.* 2013. Paths to selection on life history loci in different natural environments across the native range of Arabidopsis thaliana. *Mol Ecol.* **22**(13):3552–3566.

Franks SJ, Sim S, Weis AE. 2007. Rapid evolution of flowering time by an annual plant in response to a climate fluctuation. *Proc Natl Acad Sci U S A.* **104**(4):1278–1282.

Frenkel EM, Good BH, Desai MM. 2014. The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations. *Genetics.* **196**(4):1217.

Fulgione A, Neto C, Elfarargi AF, Tergemina E, Ansari S, Göktay M, Dinis H, Döring N, Flood PJ, Rodriguez-Pacheco S, *et al.* 2022. Parallel reduction in flowering time from de novo mutations enable evolutionary rescue in colonizing lineages. *Nat Commun.* **13**(1):1461.

Gillespie JH. 1983a. A simple stochastic gene substitution model. *Theor Popul Biol.* **23**(2):202–215.

Gillespie JH. 1983b. Some properties of finite populations experiencing strong selection and weak mutation. *Am Nat.* **121**(5):691–708.

Gomulkiewicz R, Holt RD, Barfield M, Nuismer SL. 2010. Genetics, adaptation, and invasion in harsh environments. *Evol Appl.* **3**(2):97–108.

Gomulkiewicz R, Houle D. 2009. Demographic and genetic constraints on evolution. *Am Nat.* **174**(6):E218–E229.

Good BH, McDonald MJ, Barrick JE, Lenski RE, Desai MM. 2017. The dynamics of molecular evolution over 60,000 generations. *Nature.* **551**(7678):45–50.

Gould NESJ, Eldredge N. 1972. Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM, editors. *Models in paleobiology.* San Francisco: Freeman Cooper. p. 82–115.

Hall MC, Willis JH. 2006. Divergent selection on flowering time contributes to local adaptation in Mimulus Guttatus populations. *Evolution.* **60**(12):2466–2477.

Hayward LK, Sella G. 2022. Polygenic adaptation after a sudden change in environment. *ELife.* **11**:e66697.

He F, Arce AL, Schmitz G, Koornneef M, Novikova P, Beyer A, de Meaux J. 2016. The footprint of polygenic adaptation on stress-responsive Cis-regulatory divergence in the Arabidopsis genus. *Mol Biol Evol.* **33**(8):2088–2101.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science.* **313**(5783):101–104.

Hurme P, Sillanpää MJ, Arjas E, Repo T, Savolainen O. 2000. Genetic basis of climatic adaptation in Scots pine by Bayesian quantitative trait locus analysis. *Genetics.* **156**(3):1309–1322.

Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, *et al.* 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature.* **484**(7392):55–61.

Josephs EB, Lee YW, Stinchcombe JR, Wright SI. 2015. Association mapping reveals the role of purifying selection in the maintenance of genomic variation in gene expression. *Proc Natl Acad Sci U S A.* **112**(50):15390–15395.

Josephs EB, Stinchcombe JR, Wright SI. 2017. What can genome-wide association studies tell us about the evolutionary forces maintaining genetic variation for quantitative traits? *New Phytol.* **214**(1):21–33.

Kauffman S, Levin S. 1987. Towards a general theory of adaptive walks on rugged landscapes. *J Theor Biol.* **128**(1):11–45.

Keurentjes JJB, Bentsink L, Alonso-Blanco C, Hanhart CJ, Blankestijn-De Vries H, Effgen S, Vreugdenhil D, Koornneef M. 2007. Development of a near-isogenic line population of Arabidopsis thaliana and comparison of mapping power with a recombinant inbred line population. *Genetics.* **175**(2):891–905.

Kimura M. 1983. *The neutral theory of molecular evolution.* Cambridge Core: Cambridge University Press.

Koch EM, Sunyaev SR. 2021. Maintenance of complex trait variation: classic theory and modern data. *Front Genet.* **12**:763363.

Korves TM, Schmid KJ, Caicedo AL, Mays C, Stinchcombe JR, Purugganan MD, Schmitt J. 2007. Fitness effects associated with the major flowering time gene FRIGIDA in *Arabidopsis thaliana* in the field. *Am Nat.* **169**(5):E141–E157.

Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerová M, Rubin C-J, Wang C, Zamani N, *et al.* 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature.* **518**(7539):371–375.

Lee YW, Gould BA, Stinchcombe JR. 2014. Identifying the genes underlying quantitative traits: a rationale for the QTN programme. *AoB PLANTS.* **6**:plu004.

Léran S, Varala K, Boyer J-C, Chiurazzi M, Crawford N, Daniel-Vedele F, David L, Dickstein R, Fernandez E, Forde B, *et al.* 2014. A unified nomenclature of NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family members in plants. *Trends Plant Sci.* **19**(1):5–9.

Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO. 2010. Association mapping of local climate-sensitive quantitative trait loci in Arabidopsis thaliana. *Proc Natl Acad Sci U S A.* **107**(49):21199–21204.

Liu X, Li YI, Pritchard JK. 2019. Trans effects on gene expression can drive omnigenic inheritance. *Cell.* **177**(4):1022–1034.e6.

Ludlow, M. M. (1989). Strategies of response to water stress. Structural and functional responses to environmental stresses: water shortage. 14th International Botanical Congress, Berlin, July 24-August 1, 1987.

Matuszewski S, Hermisson J, Kopp M. 2014. Fisher's geometric model with a moving optimum. *Evolution.* **68**(9):2571–2588.

McKay JK, Richards JH, Mitchell-Olds T. 2003. Genetics of drought adaptation in *Arabidopsis thaliana*: pleiotropy contributes to genetic correlations among ecological traits. *Mol Ecol.* **12**(5):1137–1151.

Mercier S, Daudin JJ. 2001. Exact distribution for the local score of one i.i.d. Random sequence. *J Comput Biol.* **8**(4):373–380.

Morris K, Jackson SP. 2010. DAY NEUTRAL FLOWERING does not act through GIGANTEA and FKF1 to regulate CONSTANS expression and flowering time. *Plant Signal Behav.* **5**(9):1105–1107.

Morris K, Thornber S, Codrai L, Richardson C, Craig A, Sadanandom A, Thomas B, Jackson S. 2010. DAY NEUTRAL FLOWERING represses CONSTANS to prevent Arabidopsis flowering early in short days. *Plant Cell.* **22**(4):1118–1128.

Mouradov A, Cremer F, Coupland G. 2002. Control of flowering time: interacting pathways as a basis for diversity. *Plant Cell.* **14**(suppl 1):S111–S130.

Nachman MW, Hoekstra HE, D'Agostino SL. 2003. The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci U S A.* **100**(9):5268–5273.

Olsen KM, Halldorsdottir SS, Stinchcombe JR, Weinig C, Schmitt J, Purugganan MD. 2004. Linkage disequilibrium mapping of Arabidopsis CRY2 flowering time AllelesSequence data from this article have been deposited with the EMBL/GenBank data libraries under accession nos. AY576055, AY576271. *Genetics.* **167**(3):1361–1369.

Olsson K, Ågren J. 2002. Latitudinal population differentiation in phenology, life history and flower morphology in the perennial herb Lythrum salicaria. *J Evol Biol.* **15**(6):983–996.

Orr HA. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution.* **52**(4):935–949.

Orr HA. 2002. The population genetics of adaptation: the adaptation of DNA sequences. *Evolution*. **56**(7):1317–1330.

Orr HA. 2005. Theories of adaptation: what they do and don't say. *Genetica*. **123**(1):3–13.

Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M. 2010. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science*. **327**(5961):92–94.

Pallares LF, Lea AJ, Han C, Filippova EV, Andolfatto P, Ayroles JF. 2023. Dietary stress remodels the genetic architecture of lifespan variation in outbred Drosophila. *Nat Genet*. **55**(1):123–129.

Peichel CL, Marques DA. 2017. The genetic and molecular architecture of phenotypic diversity in sticklebacks. *Phil Trans R Soc B: Biol Sci*. **372**(1713):20150486.

Phee B-K, Kim J-I, Shin DH, Yoo J, Park K-J, Han Y-J, Kwon Y-K, Cho M-H, Jeon J-S, Bhoo SH, et al. 2008. A novel protein phosphatase indirectly regulates phytochrome-interacting factor 3 via phytochrome. *Biochem J*. **415**(2):247–255.

Pien S, Fleury D, Mylne JS, Crevillen P, Inzé D, Avramova Z, Dean C, Grossniklaus U. 2008. ARABIDOPSIS TRITHORAX1 dynamically regulates FLOWERING LOCUS C activation via histone 3 lysine 4 trimethylation. *Plant Cell*. **20**(3):580–588.

Provine WB. 2001. *The origins of theoretical population genetics*. Chicago (IL): University of Chicago Press.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. **81**(3):559–575.

Rausher MD, Delph LF. 2015. Commentary: when does understanding phenotypic evolution require identification of the underlying genes? *Evolution*. **69**(7):1655–1664.

Reed DR, Lawler MP, Tordoff MG. 2008. Reduced body weight is a common effect of gene knockout in mice. *BMC Genet*. **9**:4.

Robertson A. 1967. The nature of quantitative genetic variation. *Heritage from Mendel*. Madison (WI): University of Wisconsin Press. p. 265–280.

Rockman MV. 2012. The QTN program and the alleles that matter for evolution: all that's Gold does not glitter. *Evolution*. **66**(1):1–17.

Rogers SM, Tamkee P, Summers B, Balabahadra S, Marks M, Kingsley DM, Schluter D. 2012. GENETIC SIGNATURE OF ADAPTIVE PEAK SHIFT IN THREESPINE STICKLEBACK. *Evolution*. **66**(8):2439–2450.

Saccheri IJ, Rousset F, Watts PC, Brakefield PM, Cook LM. 2008. Selection and gene flow on a diminishing cline of melanic peppered moths. *Proc Natl Acad Sci U S A*. **105**(42):16212–16217.

Saleh A, Alvarez-Venegas R, Yilmaz M, Le O, Hou G, Sadder M, Al-Abdallat A, Xia Y, Lu G, Ladunga I, et al. 2008. The highly similar Arabidopsis homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. *Plant Cell*. **20**(3):568–579.

Salomé PA, Bomblies K, Laitinen RAE, Yant L, Mott R, Weigel D. 2011. Genetic architecture of flowering-time variation in Arabidopsis thaliana. *Genetics*. **188**(2):421–433.

Sanagi M, Aoyama S, Kubo A, Lu Y, Sato Y, Ito S, Abe M, Mitsuda N, Ohme-Takagi M, Kiba T, et al. 2021. Low nitrogen conditions accelerate flowering by modulating the phosphorylation state of FLOWERING BHLH 4 in Arabidopsis. *Proc Natl Acad Sci U S A*. **118**(19):e2022942118.

Sanchez-Bermejo E, Zhu W, Tasset C, Eimer H, Sureshkumar S, Singh R, Sundaramoorthi V, Colling L, Balasubramanian S. 2015. Genetic architecture of natural variation in thermal responses of Arabidopsis1[OPEN]. *Plant Physiol*. **169**(1):647–659.

Schluter D, Marchinko KB, Arnegard ME, Zhang H, Brady SD, Jones FC, Bell MA, Kingsley DM. 2021. Fitness maps to a large-effect locus in introduced stickleback populations. *Proc Natl Acad Sci U S A*. **118**(3):e1914889118.

Schoustra SE, Bataillon T, Gifford DR, Kassen R. 2009. The properties of adaptive walks in evolving populations of fungus. *PLoS Biol*. **7**(11):e1000250.

Sella G, Barton NH. 2019. Thinking about the evolution of Complex traits in the era of genome-wide association studies. *Annu Rev Genomics Hum Genet*. **20**(1):461–493.

Shafiq S, Berr A, Shen W-H. 2014. Combinatorial functions of diverse histone methylations in Arabidopsis thaliana flowering time regulation. *New Phytol*. **201**(1):312–322.

Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jónsson B, Schluter D, Kingsley DM. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*. **428**(6984):717–723.

Silander OK, Tenaillon O, Chao L. 2007. Understanding the evolutionary fate of finite populations: the dynamics of mutational effects. *PLoS Biol*. **5**(4):e94.

Simons YB, Bullaughey K, Hudson RR, Sella G. 2018. A population genetic interpretation of GWAS findings for human quantitative traits. *PLoS Biol*. **16**(3):e2002985.

Simons YB, Mostafavi H, Smith CJ, Pritchard JK, Sella G. 2022. Simple scaling laws control the genetic architectures of human complex traits (p. 2022.10.04.509926). bioRxiv. https://doi.org/10.1101/2022.10.04.509926

Song YH, Estrada DA, Johnson RS, Kim SK, Lee SY, MacCoss MJ, Imaizumi T. 2014. Distinct roles of FKF1, Gigantea, and Zeitlupe proteins in the regulation of constans stability in Arabidopsis photoperiodic flowering. *Proc Natl Acad Sci U S A*. **111**(49):17672–17677.

Speidel L, Forest M, Shi S, Myers SR. 2019. A method for genome-wide genealogy estimation for thousands of samples. *Nat Genet*. **51**(9):1321–1329.

Stearns SC. 1992. *The evolution of life histories*. Oxford (UK): Oxford University Press.

Steiner CC, Weber JN, Hoekstra HE. 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biol*. **5**(9):e219.

Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS, Purugganan MD, Schmitt J. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene FRIGIDA. *Proc Natl Acad Sci U S A*. **101**(13):4712–4717.

Szendro IG, Franke J, de Visser JAGM, Krug J. 2013. Predictability of evolution depends nonmonotonically on population size. *Proc Natl Acad Sci U S A*. **110**(2):571–576.

Tessadori F, Schulkes RK, van Driel R, Fransz P. 2007. Light-regulated large-scale reorganization of chromatin during the floral transition in Arabidopsis. *Plant J*. **50**(5):848–857.

Turner TL, Stewart AD, Fields AT, Rice WR, Tarone AM. 2011. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in Drosophila melanogaster. *PLoS Genet*. **7**(3):e1001336.

van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ. 2016. The industrial melanism mutation in British peppered moths is a transposable element. *Nature*. **534**(7605):102–105.

van't Hof AE, Edmonds N, Dalíková M, Marec F, Saccheri IJ. 2011. Industrial melanism in British peppered moths has a singular and recent mutational origin. *Science*. **332**(6032):958–960.

Võsa U, Claringbould A, Westra H-J, Bonder MJ, Deelen P, Zeng B, Kirsten H, Saha A, Kreuzhuber R, Yazar S, et al. 2021. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet*. **53**(9):1300–1310.

Wallace AR. 1855. On the law which has regulated the introduction of species. *Ann Magaz Nat Hist, 2 Ser* **16**:184–196.

Walsh B, Lynch M. 2018. *Evolution and selection of quantitative traits*. Sunderland (MA): Sinauer Associates.

Wang L, Ji Y, Hu Y, Hu H, Jia X, Jiang M, Zhang X, Zhao L, Zhang Y, Jia Y, et al. 2019. The architecture of intra-organism mutation rate variation in plants. *PLoS Biol*. **17**(4):e3000191.

Weber JN, Peterson BK, Hoekstra HE. 2013. Discrete genetic modules are responsible for complex burrow evolution in peromyscus mice. *Nature*. **493**(7432):402–405.

Weng M-L, Becker C, Hildebrandt J, Neumann M, Rutter MT, Shaw RG, Weigel D, Fenster CB. 2019. Fine-Grained analysis of spontaneous mutation Spectrum and frequency in Arabidopsis thaliana. *Genetics*. **211**(2):703–714.

Yengo L, Vedantam S, Marouli E, Sidorenko J, Bartell E, Sakaue S, Graff M, Eliasen AU, Jiang Y, Raghavan S, *et al.* 2022. A saturated map of common genetic variants associated with human height. *Nature*. **610**(7933):704–712.

Zan Y, Carlborg Ö. 2018. A multilocus association analysis method integrating phenotype and expression data reveals multiple novel associations to flowering time variation in wild-collected Arabidopsis thaliana. *Mol Ecol Resources*. **18**(4):798–808.

Zan Y, Carlborg Ö. 2019. A polygenic genetic architecture of flowering time in the worldwide Arabidopsis thaliana population. *Mol Biol Evol*. **36**(1):141–154.

Zhou X, Carbonetto P, Stephens M. 2013. Polygenic modeling with Bayesian sparse linear mixed models. *PLoS Genet*. **9**(2): e1003264.

Zhou X, Stephens M. 2012. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. **44**(7): 821–824.