American Genetic Association

Journal of Heredity, 2019, 445–454
doi:10.1093/jhered/esz026
Symposium Article
Advance Access publication June 3, 2019

OXFORD

Symposium Article

Evaluating Population Genomic Candidate Genes Underlying Flowering Time in *Arabidopsis thaliana* Using T-DNA Insertion Lines

Veronica K. Chong and John R. Stinchcombe

From the Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON M5S3B2, Canada (Chong and Stinchcombe)

Address correspondence to John R. Stinchcombe at the address above, or e-mail: john.stinchcombe@utoronto.ca

Received September 11, 2018; First decision January 9, 2019; Accepted April 16, 2019.

Corresponding Editor: Anne Bronikowski

Abstract

Population genomic scans have emerged as a powerful tool to detect regions of the genome that are potential targets of selection. Despite the success of genomic scans in identifying novel lists of loci potentially underlying adaptation, few studies proceed to validate the function of these candidate genes. In this study, we used transfer-DNA (T-DNA) insertion lines to evaluate the effects of 27 candidate genes on flowering time in North American accessions of *Arabidopsis thaliana*. We compared the flowering time of T-DNA insertion lines that knock out the function of a candidate gene obtained from population genomic studies to a wild type under long- and short-day conditions. We also did the same for a collection of randomly chosen genes that had not been identified as candidates. We validated the well-known effect of long-day conditions in accelerating flowering time and found that gene disruption caused by insertional mutagenesis tends to delay flowering. Surprisingly, we found that knockouts in random genes were just as likely to produce significant phenotypic effects as knockouts in candidate genes. T-DNA insertions at a handful of candidate genes that had previously been identified as outlier loci showed significant delays in flowering time under both long and short days, suggesting that they are promising candidates for future investigation.

Keywords: adaptation, flowering time, population genomics, T-DNA

A central goal in evolutionary biology is to understand the genetic basis underlying ecologically important traits (Stinchcombe and Hoekstra 2008; Orr 2009; Hohenlohe et al. 2018). Population genomic scans have emerged as a prominent approach to studying adaptation, especially, as the ability to gather genome-wide marker data has expanded (François et al. 2016). These scans typically involve surveying genetic variability across the genome to detect "outlier" loci with unusual patterns of divergence or polymorphism (relative to genome-wide averages or statistical hypothesis tests), which is interpreted to be due to selection (Weigel and Nordborg 2005; Stinchcombe and Hoekstra 2008; de Villemereuil et al. 2014; Ahrens et al. 2018). One potential advantage of population genomic scans is that they do not require a priori knowledge of the agents of selection or even the traits that were under selection (Hohenlohe et al. 2010), which may give a more unbiased picture of adaptation. Population genomic scans, however, are not a panacea. First, they are prone to

false discoveries and numerous statistical challenges, depending on sampling designs, statistical frameworks, and population structure (see e.g., Li et al. 2008; Schoville et al. 2012; Lotterhos and Whitlock 2015; Capblancq et al. 2018). Second, many of the identified outlier loci or regions are never subjected to experimental or functional validation (François et al. 2016; Talbot et al. 2017). Without functional assays, it is difficult to determine whether variation at the detected regions affects phenotypes or fitness or can be confidently linked to adaptation (Weigel and Nordborg 2015; Barrett and Hoekstra 2011; Pavey et al. 2012). Here, we use gene knockouts in *Arabidopsis thaliana* to examine whether disrupting candidate genes identified through population genomic scans indeed affect phenotypes.

There are several potential ways to validate genes or genetic regions identified in population genomics studies, including in vitro functional assays, fine-mapping, and genetic knockouts (Ballinger and Benzer 1989; Bouché and Bouchez 2001; Skarnes et al. 2011). These varied experimental approaches have at their core some manipulation of gene function; the phenotypes of the manipulated genotypes can then be compared to that of unmanipulated wild-type controls (Bouchez and Höfte 1998). For example, Colosimo et al. (2005) validated the positive effect of *Eda,* a previously identified candidate gene, on the number of lateral plates formed in the 3-spine sticklebacks (*Gasterosteus aculeatus*) by comparing individuals that carry an *Eda* transgene to control siblings. Similarly, a study on melanism in the deer mice (*Peromyscus maniculatus*) compared the coat color of an in vitro embryonic mice and a wild type and found that melanic hair is caused by a recessive allele in the *Agouti* locus (Kingsley et al. 2009). Genetic knockouts are commonly used in studies on model organisms such as *Drosophila melanogaster*, *Caenorhabditis elegans*, and *A. thaliana*; although similar resources are currently lacking in most nonmodel organisms (Skarnes et al. 2011; Ellegren 2014). In *A. thaliana*, gene knockouts are commonly generated by inserting agrobacterium-mediated transfer-DNA (T-DNA) into specific regions in the genome to disrupt gene function (Krysan et al. 1999). T-DNA insertions are chemically and physically stable over multiple generations, do not travel to other locations in the genome (Martienssen 1998; Østergaard and Yanofsky 2004), and often cause large, detectable effects on gene function (Krysan et al. 1999). Recent developments in CRISPR/CAS9 technology suggest further possibilities for validating the effects of candidate regions or loci (see Turner 2014), although to our knowledge this has not yet been applied to candidates identified by population genomic studies.

The ideal genetic evidence to determine if outlier loci are indeed responsible for differentiation would be to use some sort of genetic modification (repeated backcrossing, transgenics, CRISPR, etc.) to move naturally occurring alleles into a isogenic test background. In this manner, the effects of a single allele on the phenotype can be verified. For many population genomics studies, however, this is likely to remain challenging. Most population genomic scans will obtain dozens to 100s of outlier loci; some method of prioritizing or triaging of candidate loci or regions will be required. We suggest that use of knockouts such as T-DNA insertions is a useful screening tool to identify whether a gene, *in principle*, can affect the phenotype of interest. That is, a knockout or T-DNA insertion can allow an investigator to ask "does the presence or absence of this gene affect my phenotype of interest" rather than "does the observed, naturally occurring variant affect the phenotype of interest?" Given the risk of false positives in population genomic scans, such a screening tool can be useful for identifying promising genes for future study about naturally occurring variants. The approach implicitly assumes that knocking out a gene entirely is more likely to have an effect on a phenotype than many naturally occurring mutations (nonsynonymous mutations, frameshifts, premature start/stop codons, etc.), such that the absence of an effect of a whole-gene knockout probably means that detailed genetic investigation of naturally occurring variants at these genes is a lower priority.

Flowering is one of the most important plant life-history traits, as it determines a plant's reproductive investment and ultimately its fitness (Srikanth and Schmid 2011). Flowering marks the transition from vegetative to reproductive phase of a plant, and thus the timing of flowering directly determines whether seed set is completed under favorable conditions (Andrés and Coupland 2012). The mouse-ear cress (*Arabidopsis thaliana*) is a popular model organism for exploring the timing of flowering in angiosperms (Koornneef et al. 2004). Consistent findings in the literature suggest that *A. thaliana* populations exhibit spatial clines in flowering time and that abiotic factors such as daylength, precipitation, and temperature contribute significantly to these clinal differences (Stinchcombe et al. 2004; Cookson et al. 2007; Amasino and Michaels 2010; Munguía-Rosas et al. 2011). Natural variation in flowering time of different populations of *A. thaliana* suggests that flowering time is adapted to various environmental conditions (Amasino 2010; Srikanth and Schmid 2011; Andrés and Coupland 2012). The genetic basis of flowering time has also been intensively studied in *A. thaliana* (see Bloomer and Dean 2017 for a review). The effects of major genes that regulate flowering time (such as *FLOWERING L C* [*FLC*], *FRIGIDA* [*FRI*], *PHYC, MAF2,* etc.) as well as their expression in response to seasonal environmental cues have been extensively studied (Geraldo et al. 2009; Kim et al. 2009). Two major effect genes, *FLC* and *FRI*, are known to interact epistatically and explain up to 70% of flowering time variation in some accessions of *A. thaliana* (Caicedo et al. 2004; Shindo et al. 2005). Interestingly, polymorphism at *FLC* itself only explains a portion of its effects on flowering time, while its expression is significantly associated with flowering time: these data suggest that other loci influence expression of *FLC* and that expression influences flowering time (Sasaki et al. 2018). However, flowering time is a complex quantitative trait, and it is likely that there are many undiscovered genes with small effects that are also acting on this trait (Michaels 2009). It remains unknown how these undiscovered genes interact with environmental signals and other genes and what roles they play in molecular pathways leading to flowering (Adrian et al. 2009; Amasino 2010).

Using population genomic scans, Gould and Stinchcombe (2017) discovered novel candidate genes that are putatively linked to flowering time differentiation in the introduced range of *A. thaliana*. They tested for outlier loci that were highly differentiated between early- and late-flowering genotypes. Gould and Stinchcombe (2017) detected outlier single nucleotide polymorphisms (SNPs) or variants showing high X^TX scores. X^TX is a F_{ST}-like statistic measuring population differentiation, with a few key differences (see Gunther and Coop 2013). First, it accounts for variation and covariation in allele frequencies among populations, and uneven sampling among groups, which gives it greater power than traditional F_{ST} (Gunther and Coop 2013). The expected value for X^TX is the number of groups being compared (i.e., 2 in the case of early- vs. late-flowering groups). Gould and Stinchcombe (2017) estimated the variance and covariance in allele frequencies across populations using ~3000 synonymous SNPs. They then compared X^TX scores for high-impact variants (start/stop codons, frameshift mutations, etc.) to the distribution of X^TX values for ~20 000 synonymous SNPs. Gould and Stinchcombe (2017) identified candidate genes as those containing

SNPs with high X^TX scores and those falling in the extremes of genome-wide distributions of either π or the cross-population composite likelihood ratio (XP-CLR) test statistic. The outlier variants they detected were not enriched for a priori known or suspected candidate genes and are thus potentially novel loci. Despite the steps taken by Gould and Stinchcombe (2017) to guard against false positives (conditioning on covariation in allele frequencies among populations, comparison of high-impact and synonymous SNP X^TX scores, requiring other population genetic evidence of a selective sweep), an important next step is to experimentally validate that variation at these genes can, in principle, affect flowering time.

In this study, we used T-DNA insertion lines to evaluate the phenotypic effects of gene disruption in 27 of the candidate genes for flowering time in North American accessions of *A. thaliana* identified by Gould and Stinchcombe (2017). Because flowering is a quantitative trait that integrates multiple environmental and genetic signals and across growth and development from germination to the onset of flowering, we also selected 27 T-DNA insertion lines chosen at random to serve as a point of comparison for the overall effects of T-DNA insertion. We compared the phenotype of each T-DNA insertion line to the wild type to infer the effects of targeted gene disruption on flowering time under both short and long-day conditions. Collectively, we found that T-DNA insertions tend to delay flowering relative to the wild type and that the phenotypic effects of T-DNA insertions are correlated across daylength environments. While we failed to detect differences in the mean effect size of knocking out candidate genes versus genes chosen at random, we were able to validate that T-DNA insertion has strong, consistent, and significant effects on flowering time for a handful of candidate and randomly chosen genes.

Methods

T-DNA Insertion Lines

The SALK Institute Genome Analysis Laboratory (SIGnAL) generated T-DNA insertion lines in the Columbia background (Col-0) for approximately 75% of the *A. thaliana* genome (Alonso et al. 2003). Some T-DNA insertion collections (e.g., GABI-Kat, SAIL, WiscDsLox) generated by other *Arabidopsis* knockout facilities also cover parts of the remaining 25% of the genome (Sessions et al. 2002; Woody et al. 2007; Kleinboelting et al. 2012). The T-DNA lines are sometimes referred to as a "unimutant" collection, although many contain more than one insertion (O'Malley and Ecker 2010; Valentine et al. 2012). While investigations of the phenotypic effects of multiple insertions are rare, Rutter et al. (2017) found no relationship between insertion number and fitness (fruit number) in a sample of 113 T-DNA insertion lines. We attempted to mitigate the effects of lines containing multiple insertions by choosing GABI-Kat lines guaranteed to have a single insertion and prioritizing T-DNA lines that had been verified by next-generation sequencing. We obtained T-DNA insertion lines from the Arabidopsis Biological Resource Center and the Nottingham Arabidopsis Stock Centre; preliminary analyses indicated no significant differences between types of T-DNA insertion lines (e.g., GABI-Kat vs. others). For 27 of the candidate genes, we selected one T-DNA insertion line that targeted exonic regions in the Col-0 background and was homozygous at the targeted region. Based on these criteria, we also selected a set of T-DNA lines at random sites across the genome that matched the chromosomal distribution of candidate genes. We consider consistent differences between a T-DNA line and the wild-type control in both treatments as evidence that a gene potentially warrants further study as a candidate. We use the T-DNA insertions in random lines as a point of comparison to determine whether insertions at candidates identified through differentiation and patterns of nucleotide diversity differ in any obvious way from a random sample in their phenotypic effect size, consistency across environments, or direction of effects.

We verified that the T-DNA insertion lines were homozygous at the insertion site with polymerase chain reaction (PCR) and agarose gel electrophoresis analyses. We first stratified seeds in 0.15 mg/100 mL agarose solution at 4 °C for 5 days and then germinated seeds in soil in a growth chamber at 16-h day cycles (22/20 °C day/night temperature). We extracted genomic DNA from leaves of 4-week-old seedlings using Qiagen DNeasy Plant Mini Kits (Qiagen Inc., Toronto, Canada). We then performed PCR using target-specific primers designed by the SIGnaL and protocol developed by Joly-Lopez et al. (2016); we visualized reaction products on an agarose gel to determine the genotype of each insertion line. We gathered selfed seeds from 27 candidates and 27 random lines verified to be homozygous at the insertion site for experimental work (see Supplementary Table S1) as well as selfed seeds from the wild-type control (CS70000) under the same long-day environmental conditions.

Common Garden Experiment

We planted seeds set by the 54 verified T-DNA insertion lines alongside the Col-0 ecotype (which served as the wild-type control) in 4 programmable growth chambers. Because daylength is known to affect flowering time in *A. thaliana* and the Samis et al. (2012) and Gould and Stinchcombe (2017) studies of North American *A. thaliana* used flowering time from outdoor experiments with naturally changing daylengths, we included 2 daylength treatments in our common garden experiment. We used 2 long-day chambers (16/8-h light/dark cycle; 22/20 °C day/night temperature) and 2 short-day chambers (8/16-h light/dark cycle; 22/20 °C day/night temperature). We replicated chambers to ensure that daylength effects were not confounded with chamber effects.

To synchronize germination, we stratified seeds in 0.15 mg/100 mL agarose solution for 7 days at 4 °C. We then planted each seed into a standard conetainer with saturated Sunshine Mix #1 soil (Sun Gro Horticulture, Agawam, MA). In each chamber, we grew 8 replicates for each of the 55 genetic lines in a randomized block design, yielding 440 plants per chamber (N = 1760 experimental plants in the 4 chambers); blocks corresponded to shelves on the chamber. The experiment spanned a 4-month period, during which we watered plants every 5 days by saturating soil in standing trays of water for approximately 3 h. Throughout the experiment, we monitored the plants regularly, recording bolting date (visible differentiation of the apical meristem) and flowering date. We measured flowering time as the number of elapsed days between germination and flowering date.

Data Analysis

Plastic Responses to Daylength Treatment

We first assessed the effects of daylength treatment, genotype, their interaction, and growth chamber on flowering time using the following nested, mixed-model analysis of variance (Anova; lme4, R):

$$\log_{10}(\text{flowering time}) = \text{daylength} + \text{genotype} + \text{daylength} \times \text{genotype} + \text{chamber}(\text{treatment}) \quad (1)$$

where flowering time ($\log_{10}$-transformed) was modeled as the dependent variable, with fixed effects of daylength treatment, genotype (i.e., the T-DNA lines and the control), and their interaction; growth chamber nested within daylength treatment was included as a random effect. In this model, the daylength effect tests for plastic responses to the treatment, the genotype effect tests for genetic variation in flowering time among all the lines used, and the daylength × genotype term tests for genetic variation among the T-DNA lines in their plastic flowering response to daylength. We designated line as a fixed effect because the candidate genes and hence T-DNA insertion lines were chosen a priori.

We took several steps to satisfy the assumptions of an Anova-based model. First, we excluded outliers that were greater than 4 standard deviations from the mean in each treatment and thus eliminated 12 observations (0.7%) from the data set (compared to an expectation of 0.006% of the data under normality). These 12 outliers came from 10 T-DNA insertion lines each with 1 or 2 replicates, indicating that this procedure did not lead to differential elimination of T-DNA lines. We nevertheless verified that removing the outliers did not significantly affect the results of our hypothesis tests. We analyzed $\log_{10}$-transformed data for all hypothesis tests (both flowering time and log10-transformed flowering time were significantly different from normal), although we present results and figures on the original scale for clarity. We verified the results of hypothesis tests for the model above with randomization tests using procedures developed by Cassell (2002) in SAS software. In short, we randomly reshuffled flowering time (without replacement), assigned it to dependent variables, and fit model (1) with Proc Mixed (SAS v. 9.4); we repeated this reshuffling and model fitting 1000 times to estimate how often we would obtain results by chance alone (i.e., when there is only a randomly assigned relationship between flowering time and the independent variables). Results from parametric tests matched those using randomization and thus we present the former.

Comparing T-DNA Insertion Lines to Wild-Type Control

After verifying that there were significant treatment and treatment by line interactions, we performed follow-up analyses within each daylength treatment to examine whether the T-DNA insertion lines differed significantly from wild-type controls. We included chamber, genetic line, and their interaction in the following fixed effects Anova for flowering time:

$$\log_{10}(\text{flowering time}) = \text{chamber} + \text{line} + \text{chamber} \times \text{line} \quad (2)$$

For these analyses, we initially included the type of T-DNA insertion line (candidate or random) as a fixed effect nested within genetic line to test whether there was an overall difference between randomly selected lines and candidate lines within a daylength treatment. We failed to detect significant effects of the type of T-DNA insertion line on flowering time and subsequently dropped this term from our models. From model (2), we tested whether each T-DNA insertion line differed from the wild type using a Dunnett's test. The Dunnett's test controls the maximum experiment-wise error rate at the specified level of α (0.05), which prevents the inflation of Type-I error from multiple comparisons and thus the need to control for multiple comparisons using approaches like Bonferroni corrections. The Dunnett's test is designed specifically for comparing the mean of one group against a control (i.e., T-DNA lines vs. wild type), rather than all groups against each other (Dunnett 1955; see Moyle and Graham 2005 for an example in the context of genetics). Because we had equal sampling of wild type and T-DNA lines, rather than allocating more replicates to the wild type, we expect these Dunnett's tests to be conservative.

Comparison of T-DNA Effects and Population Genetic Differentiation

We next evaluated the similarity between the effects of T-DNA insertion in the 2 environments and whether there was any relationship between the phenotypic effects and the population genetic evidence that led a gene to be identified as a candidate. We compared the phenotypic effects of a T-DNA insertion under short days to that under long days using the deviation between the T-DNA insertion line and the control as the phenotype in each environment. For this analysis, we examined the correlation between the effect size of a T-DNA insertion under long days with its effects under short days.

Second, we evaluated whether there is a correlation between the phenotypic effects of the T-DNA insertion lines and the X^TX scores of their corresponding genes reported by Gould and Stinchcombe (2017). For the candidate genes, we used the X^TX scores in Gould and Stinchcombe (2017), which were used to identify the genes as candidates; for the random genes, we calculated X^TX scores from Gould and Stinchcombe's (2017) data. In the case of genes having multiple outlier variants, we selected the variant with the highest X^TX score following Gould and Stinchcombe (2017) for consistency. In total, X^TX scores were available for 48 of the 54 phenotyped T-DNA lines. Missing X^TX scores were either the result of the absence of outlier variants in the loci or large deletions that hindered the detection of outlier variants in the region (Gould and Stinchcombe 2017).

Results

Plasticity and G×E for Flowering Time

As is well described for *A. thaliana* (e.g., Cookson et al. 2007; Andrés and Coupland 2012), we found a plastic response to long days of accelerated flowering. Plants under the long-day treatment flowered significantly earlier than those in the short-day treatment (mean ± standard error: long day 30.04 ± 0.07 days, short day 72.44 ± 0.18 days; $P < 0.001$; Table 1). The Anova model also showed that there were significant genetic line and genetic line by treatment interaction (both $P < 0.0001$; Table 1), indicating genetic variation in flowering and in the plastic response to daylength.

We subsequently analyzed the effects of T-DNA insertion lines on flowering time separately for each daylength treatment. We again found significant genetic line effects, as well as significant chamber effects in both the long- and short-day treatments (all $P < 0.0001$; Table 2). However, we did not detect any line by chamber interaction in either treatment (both $P > 0.2$; Table 2). The significant chamber term suggests subtle microenvironmental differences (e.g., light intensity, humidity, precision of temperature control) between the 2 chambers for each treatment, but these effects did not differ across genetic lines (due to the absence of a line by chamber interaction).

Table 1. Mixed-model Anova for flowering time (days), including the effects of treatment, genetic line, and treatment by genetic line interaction on flowering time. A random effect of chamber nested within treatment was included to account for chamber effects but is not presented

Model term	F	P
Treatment	$F_{1,2} = 2033.02$	$P = 0.0005$
Line	$F_{54,1538} = 8.63$	$P < 0.0001$
Line × treatment	$F_{54,1538} = 2.85$	$P < 0.0001$

Candidate and Random T-DNA Insertion Lines Have Similar Effects

We first tested whether candidate and random T-DNA insertion lines differed in their effects on flowering time with the expectation that candidate lines would have larger phenotypic effects relative to lines that were selected at random loci. To our surprise, the analyses revealed that the type of T-DNA insertion line (candidate or randomly selected) had no significant effects on flowering time in models including both treatments or separated by treatment (all $P > 0.1$).

Table 2. Anova for flowering time in the 2 separate treatments

	Long-day treatment		Short-day treatment	
Model term	*F*	*P*	*F*	*P*
Chamber	$F_{1,765} = 37.82$	$P < 0.0001$	$F_{1,665} = 53.17$	$P < 0.0001$
Line	$F_{54,765} = 5.91$	$P < 0.0001$	$F_{54,665} = 5.26$	$P < 0.0001$
Chamber × line	$F_{54,765} = 1.15$	$P = 0.22$	$F_{54,665} = 0.99$	$P = 0.49$

T-DNA Insertion Lines Delay Flowering Time Under Both Daylengths

The majority of T-DNA insertion lines delayed flowering relative to the wild type, and this was consistent in both the long- and short-day treatments (Figure 1). In the long-day treatment, 24 insertion lines (44%, 13 candidate and 11 random) significantly delayed flowering time, while in the short-day treatment, 7 insertion lines (13%, 4 candidate and 3 random) flowered significantly later than wild type (Figure 1; Supplementary Table S2).

We identified 3 candidate and 2 random T-DNA insertion lines with substantial effects on flowering time phenotype relative to the wild type (candidate genes: AT2G43190, AT4G26095, and AT5G59930; random genes: AT1G17860 and AT2G25870; Supplementary Table S2). The 5 T-DNA insertion lines significantly delayed the timing of flowering in both treatments (Supplementary Table S2). The phenotypic effects of these lines range from 2.0 to 5.6 days (7% to 19%) in the long-day treatment and 5.4 to 13.6 days (8% to 19%) in the short-day treatment (Supplementary Table S2).

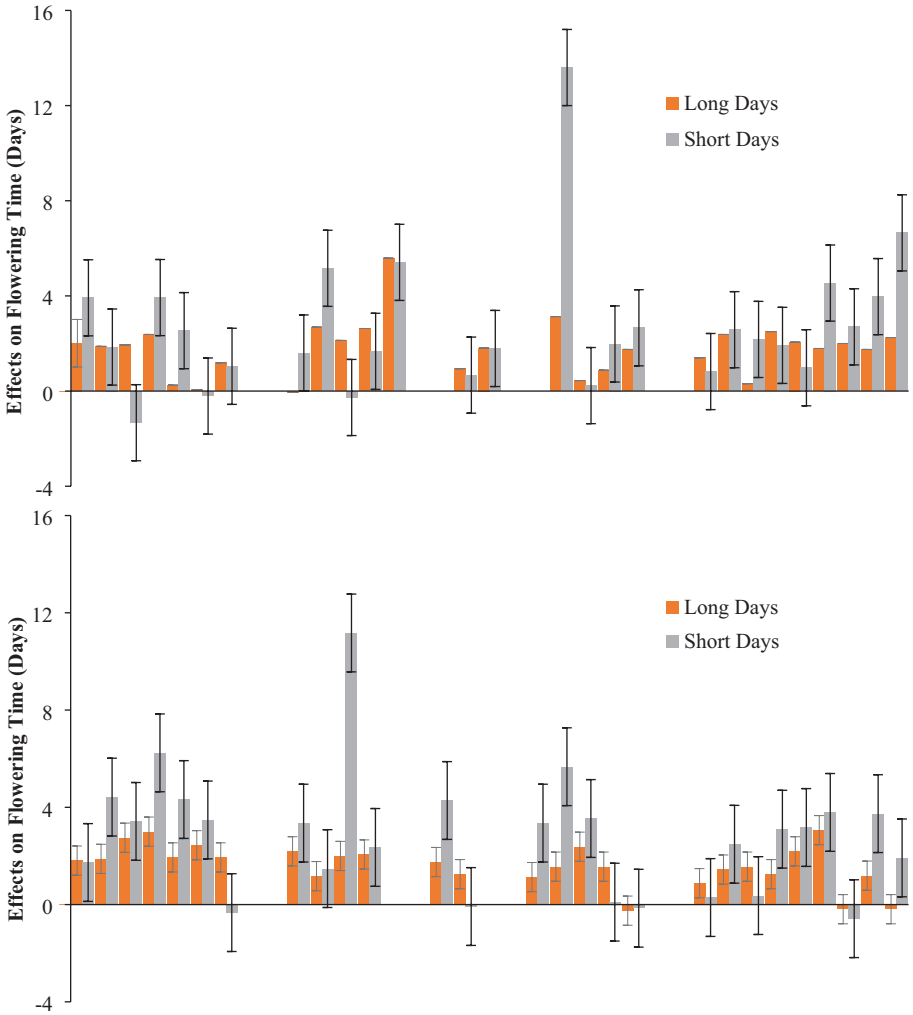

Figure 1. Bar graphs illustrating the phenotypic effects of candidate (top) and random (bottom) T-DNA insertion lines on flowering time relative to the wild type in the long- and short-day treatments. Positive values represent delay in flowering time, while negative values represent acceleration. T-DNA insertion lines are ordered by their position in the genome (left to right: chromosome 1 to chromosome 5; also see Supplementary Table S1). Error bars represent standard error.

Correlated Effects of T-DNA Insertions Across Environments

We found that the phenotypic effects of gene disruption using T-DNA insertion lines were correlated between long- and short-day conditions—T-DNA insertion lines with large effects on flowering time under long-day conditions also had large effects under short-day conditions (Figure 2). Candidate and random T-DNA insertion lines showed similar patterns (candidate: r_g = +0.49, P < 0.01; random: r_g = +0.49, P < 0.01). Many T-DNA insertion lines (65%) had larger effects on flowering time in the short-day treatment relative to that in the long-day treatment (Supplementary Table S2). There was much higher variance in flowering time of *A. thaliana* growing under short days, whereas those grown under long-day conditions exhibited more synchronized flowering time.

Genetic Differentiation is Unrelated to Phenotypic Effect Size

We next evaluated whether there is a correlation between the calculated X^TX scores reported by Gould and Stinchcombe (2017) and the phenotypic effects of their corresponding T-DNA insertion lines. We detected weak rank correlations between X^TX scores and phenotypic effects of T-DNA insertion lines in both the long- and short-day treatments (Figure 3; Supplementary Table S3). Although candidate genes had much higher X^TX scores than random genes (which is how they were defined as candidate genes in the first place), we found similar relationships between differentiation at a region (X^TX score) and the phenotypic effect of knocking it out for both types of loci. The rank correlations between X^TX values and phenotypic effects were low and nonsignificant (Figure 3; candidate in long days: r = −0.02, P = 0.9; candidate in short days: r = +0.16, P = 0.4; random in long days: r = −0.04, P = 0.9; random in short days: r = +0.19, P = 0.4). Neither the pattern nor the significance of these results changed when we used Pearson correlations. As an alternative means of investigating this, we also searched the *A. thaliana* GWAS database (https://aragwas.1001genomes.org/#/) to determine if any of the outliers identified by Gould and Stinchcombe (2017) were identified to be associated with flowering time in (mainly) European samples. We failed to find any GWAS hits, which is perhaps unsurprising given the number of unique SNPs Gould and Stinchcombe (2017) identified.

Discussion

Our experiment knocking out candidate genes for flowering time in *A. thaliana* using T-DNA insertion lines revealed 3 major results. First, gene disruption through T-DNA insertions almost universally delayed flowering time relative to the wild type. Second, we failed to detect significant differences between the phenotypic effects of candidate and random T-DNA insertion lines on flowering time or a relationship between the phenotypic effects of gene disruption and the level of population genetic differentiation at mutations in these genes that led to their identification as candidates. Finally, we identified several T-DNA insertion lines that significantly delayed flowering time under both long- and short-day conditions. The genes disrupted by these T-DNA insertion lines are excellent targets for future studies on the genetic basis of flowering time adaptation in North American accessions of *A. thaliana*. We discuss these results in turn.

T-DNA Insertion Lines Tend to Delay Flowering Time

The majority of T-DNA insertion lines—both candidate and random lines—caused delays in flowering time relative to the wild type in the common garden experiment. Results from the Dunnett's tests revealed that 24 of the 54 T-DNA insertion lines (44%) in the long-day treatment as well as 7 lines (13%) in the short-day treatment significantly affected flowering time compared to the wild-type control, all of which caused delays in flowering time phenotype. None of the T-DNA insertion lines flowered significantly earlier than the wild type under either daylength treatments. The extent of delays exhibited by T-DNA insertion lines was as large as 5.6 days (+19%) under long-day conditions and 13.6 days (+19%) under short-day

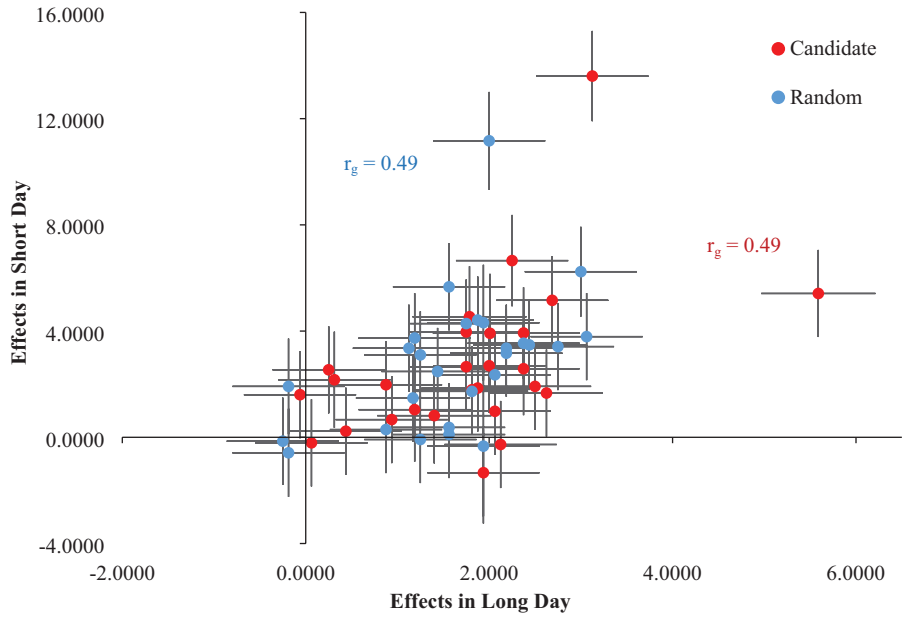

Figure 2. Scatter plot illustrating the phenotypic effects of T-DNA insertion lines on flowering time in the long- versus short-day treatment. Red dots represent line means for candidate T-DNA insertion lines; blue dots represent line means for random T-DNA insertion lines. Error bars represent standard error in each daylength treatment, while r_g-values represent correlation coefficient of candidate or random T-DNA insertion lines.

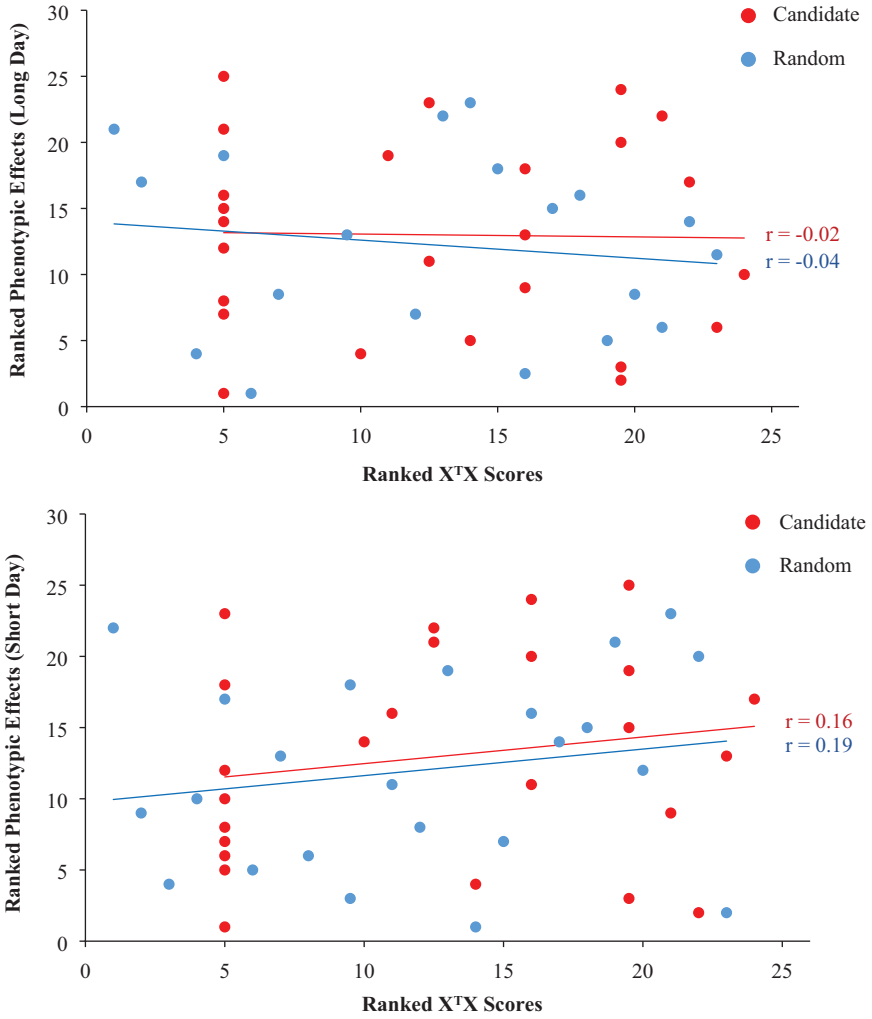

Figure 3. Correlation between X^TX scores calculated by Gould and Stinchcombe (2017) and phenotypic effects of T-DNA insertion lines on flowering time in the long- (top) and short-day (bottom) treatments. Red dots represent line means for candidate T-DNA insertion lines; blue dots represent line means for random T-DNA insertion lines.

conditions. What can explain this tendency for T-DNA insertion to delay flowering time and how does this compare to past work?

One possible explanation of the direction of our results is a combination of genetic background and growth conditions. All T-DNA insertion lines selected for this study were generated in the ecotype Col-0, which is known to be an early flowering accession (Kim et al. 2004). Many naturally occurring accessions of *A. thaliana* are late flowering due to their requirement for a prolonged period of cold (known as vernalization) to initiate floral transitions (Henderson and Dean 2004). Flowering prior to vernalization is inhibited by synergistic interactions of the loci *FLC* and *FRI* (Amasino and Michaels 2010). However, rapid cycling accessions of *A. thaliana*, such as Col-0, carry loss-of-function mutations in *FRI*, which then reduces the expression of *FLC*, thus leading to early flowering even in the absence of vernalization (Michaels and Amasino 1999; Ding et al. 2013). One consequence of this is that screens of mutant phenotypes are more challenging in early flowering backgrounds like Col-0, although most genetic resources such as T-DNAs are in this background. In an early flowering background, mutations or gene disruptions that affect flowering time may be more likely to lead to delays than acceleration. In similar fashion, our growth conditions may have led to earlier flowering (even under short days)—compared to over-wintering plants that might experience much express much longer lifespans under ecologically realistic conditions. Under ecological conditions leading to longer life spans, some of the gene disruptions that we studied could lead to accelerated flowering. Testing these hypotheses would require first, screening T-DNAs or other gene disruptions in a late-flowering background, or second, using ecologically realistic planting conditions likely to lead to over-wintering and longer life spans.

The closest point of comparison in the literature to our experiment is a set of studies by Rutter and colleagues (Rutter et al. 2010, 2012; Valentine et al. 2012; Roles et al. 2016; Rutter et al. 2017), although their focus was on how mutations affected fitness components rather than flowering time. Across a set of mutation accumulation (MA) studies, they found significant and appreciable G×E for fitness components (Rutter et al. 2012; Roles et al. 2016), which is qualitatively similar to our finding of G×E for flowering time, although the mutations produced by MA procedures are likely to be radically different than T-DNA insertions. In their MA studies, they found that mutations were likely to both increase or decrease fitness (Rutter et al. 2010), which is in contrast to our finding of T-DNAs

predominantly delaying flowering. Interestingly, they note that this could be an artifact of MA lines being derived from the early flowering ecotype, Col-0, that is adapted to lab rather than field conditions. In the most directly comparable studies, they found that between 10% and 40% of T-DNA mutations altered fitness compared to the wild type (Rutter et al. 2017), which is similar to the range we find for flowering time. Like us, they also found G×E for fitness in T-DNA lines, along with a generally positive correlation across environments (Rutter et al. 2017).

T-DNA Insertions in Candidate and Random Genes Show Similar Effects

We included a set of random T-DNA insertion lines in the experiment to control for the effects of gene disruption through insertional mutagenesis. We expected candidate T-DNA insertion lines to have larger phenotypic effects on flowering time compared to that of the random lines. However, Anova results suggested that the type of T-DNA insertion line (candidate or random) has no effect on flowering time phenotype. In fact, candidate and random T-DNA insertion lines behaved similarly under both long- and short-day conditions, with strikingly similar degree of phenotypic effects on flowering time. The lack of significant difference between candidate and random gene disruption implies that the effects of gene disruption on outlier loci were similar to that observed simply by chance from any gene disruption in the genome. Moreover, previously calculated X^TX scores for the candidate and random genes did not seem to be correlated with the phenotypic effects of their respective T-DNA insertion lines. We expected T-DNA insertion lines corresponding to genes with relatively higher X^TX scores to have larger phenotypic effects on flowering time. The similar phenotypic patterns shown by candidate and random T-DNA insertion lines as well as the mismatch between X^TX scores and phenotypic effects of knockouts at the genes could be explained by the possible false discovery of candidate genes, complex nature of flowering time adaptation, or the potential limitation of gene knockouts in mimicking natural mutations.

One possible explanation to the lack of differences between candidate and random T-DNA insertion lines is that the candidate genes identified by Gould and Stinchcombe (2017) were false positives and thus do not represent true targets of selection on flowering time. Gould and Stinchcombe (2017) pointed out that the population genomic tests used to identify the 27 candidate genes we studied here might not have much power due to its small sample size (see Gould and Stinchcombe 2017). If the set of candidate genes were not associated with flowering time adaptation, it would be reasonable to detect no difference between their effects on the phenotype and a set of genes selected randomly across the genome. Nevertheless, the use of multiple test statistics by Gould and Stinchcombe (2017) as well as their comparison of X^TX scores to the genome-wide distribution obtained from synonymous SNPs should reduce the rate of false positives (Lotterhos and Whitlock 2015). Therefore, it is worth considering whether there are other reasons behind the parallel phenotypic effects of candidate and random genes.

Another potential explanation is that flowering time is an extremely complex quantitative trait, and hence many mutations could potentially lead to changes in the phenotype (Mackay et al. 2009). Flowering time defines an angiosperm's shift from a vegetative to reproductive stage. Therefore, the timing of flowering integrates all life-history events that happens in the life cycle before reproduction (Bergelson and Roux 2010). For instance, Samis et al. (2012) found positive correlations between the number of rosette leaves, rosette diameter, and flowering time in both the introduced and native ranges of *A. thaliana*. Similarly, Debieu et al. (2013) found strong patterns of covariation among 3 major life-history traits: seed dormancy, vegetative growth rate, and flowering time in *A. thaliana* along a latitudinal cline. They also found complex genetic architectures underlying the correlated life-history traits. These correlations (if they are due to pleiotropy or tight linkage disequilibrium) suggest that variation in a single locus could have cascading effects on the phenotype of all correlated traits. Hence, mutations at loci that affect any correlated traits could potentially affect the outcome of flowering time in a cumulative manner (Salomé et al. 2011). This could explain why many loss-of-function mutations at randomly selected genes caused a significant delay in flowering time relative to the wild type in the experiment and why random genes seemed similar to candidate genes in their degree of control over flowering time phenotype.

It is also possible that while the random T-DNA insertions may affect flowering in the lab, mutations in these genes might be selected against in nature. Our screen focused primarily on flowering time, and as a consequence we were not able to evaluate other quantitative traits that might affect fitness or fitness components such as fruit and seed number, seed dormancy and germination, performance in the seed bank, or early establishment under natural conditions. It may be that mutations at the random genes targeted by T-DNAs in our experiment can in principle have an effect on phenotypes but that any such mutations would be strongly selected against under natural conditions. Testing this hypothesis would require measuring the fitness consequences of random T-DNA insertions, much like Valentine et al. (2012) and Rutter et al. (2017).

T-DNA insertion lines may also not be the best at mimicking natural mutations detected in natural accessions or populations of *A. thaliana*. T-DNA insertion lines cause a loss-of-function mutation by introducing a large T-DNA fragment (~1 kb) in the targeted locus (O'Malley et al. 2015). However, outlier variants found by Gould and Stinchcombe (2017) represent various forms of mutations such as SNPs, indels causing frameshifts or gain/loss of start/stop codons, large deletions that are up to ~9 kb, and so forth (see Supplementary Table S2 in Gould and Stinchcombe 2017). While gene knockouts may be powerful for studying natural mutations that cause the inactivation of gene function (e.g., frameshifts, large indels), they may not best reproduce the phenotypic effects of smaller-scale nonsynonymous mutations (e.g., SNPs, small indels) that result in changes in amino acid sequences rather than complete silencing of the gene (Matus et al. 2014). Thus, while 1-kb knockouts may have similar effects on phenotypes if placed in candidate genes or random genes, mutants that are more similar to those observed in nature may indeed show a difference. Testing this hypothesis would require the ability to perform mutagenesis that would more accurately mimic the natural spectrum of mutations or alternative genetic approaches such as transformation.

Future Directions: Following Up on Promising Genes

Despite potential limitations in T-DNA insertion lines in mimicking natural nonsynonymous mutations, the loss-of-function mutations at several genes that we studied have substantial effects on flowering time under both treatments, which suggests a potential relationship between these genes and flowering time. Three candidate and 2 random T-DNA insertion lines caused significant delays in flowering time compared to the wild type under both long- and

short-day conditions (candidate genes: AT2G43190, AT4G26095 and AT5G59930; random genes: AT1G17860 and AT2G25870). The 3 candidate genes include AT2G43190, which encodes for a protein involved in ribosomal RNA processing; while AT4G26095 is a potential natural antisense gene and AT5G59930 encodes for a Cysteine/Histidine-rich C1 domain family protein. The X^TX scores of AT4G26095 and AT5G59930 were approximately in the 80th and 70th percentiles for all outlier variants detected by Gould and Stinchcombe (2017), further suggesting these genes as a potential target of selection. The list of random genes consists of AT1G17860, a gene that stops or reduces the activity of an endopeptidase, and AT2G25870, a gene that controls rRNA processing in the chloroplast (information from https://www.arabidopsis.org/).

Furthermore, transgenic experiments can be conducted utilizing the alleles associated with early- and late-flowering time detected by Gould and Stinchcombe (2017). For instance, early flowering alleles could be transgenically introduced into a late-flowering accession and vice-versa. Moreover, these early- or late-flowering alleles could be separately introduced into the ecotype Col-0 to better understand their functions on flowering time in an identical genetic background. The available genomic database and genetic tools will facilitate our understanding of the genetic basis of flowering time adaptation in the introduced range of *A. thaliana*.

Supplementary Material

Supplementary material is available at *Journal of Heredity* online.

Funding

This work was supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant to J.R.S. and a graduate fellowship to V.K.C.

Acknowledgements

We thank Stephen Wright, Asher Cutter, Marla Sokolowski, and two anonymous reviewers for comments on previous versions of this manuscript. We thank Tia Harrison for help in the lab, Corlett Wood for data analysis suggestions, and Caroline Biel, Bonnie Pilkington, and Amalia Caballero for plant help.

Data Availability

We have deposited the raw data for this paper in Dryad (https://doi.org/10.5061/dryad.k5r20q0).

References

Adrian J, Torti S, Turck F. 2009. From decision to commitment: the molecular memory of flowering. *Mol Plant*. 2:628–642.

Ahrens CW, Rymer PD, Stow A, Bragg J, Dillon S, Umbers KDL, Dudaniec RY. Forthcoming 2018. The search for loci under selection: trends, biases, and progress. *Mol Ecol* 27: 1342-1356. doi:10.1111/mec.14549

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, *et al*. 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*. 301:653–657.

Amasino R. 2010. Seasonal and developmental timing of flowering. *Plant J*. 61:1001–1013.

Amasino RM, Michaels SD. 2010. The timing of flowering. *Plant Physiol*. 154:516–520.

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

Ballinger DG, Benzer S. 1989. Targeted gene mutations in *Drosophila*. *Proc Natl Acad Sci USA*. 86:9402–9406.

Barrett RD, Hoekstra HE. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet*. 12:767–780.

Bergelson J, Roux F. 2010. Towards identifying genes underlying ecologically relevant traits in Arabidopsis thaliana. *Nat Rev Genet*. 11:867–879.

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*. 68:5439–5452.

Bouché N, Bouchez D. 2001. *Arabidopsis* gene knockout: phenotypes wanted. *Curr Opin Plant Biol*. 4:111–117.

Bouchez D, Höfte H. 1998. Functional genomics in plants. *Plant Physiol*. 118:725–732.

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 USA*. 101:15670–15675.

Capblancq T, Luu K, Blum MGB, Bazin E. Forthcoming 2018. Evaluation of redundancy analysis to identify signatures of local adaptation. *Mol Ecol Resour*. 18:1223-1233 doi:10.1111/1755-0998.12906

Cassell DL. 2002. A randomized-test wrapper for SAS PROCs. In: Proceedings of the 27th Annual SAS Users Group International Conference (SUGI 27). Cary (NC): SAS Institute. p. 1–4.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G Jr, 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*. 307:1928–1933.

Cookson SJ, Chenu K, Granier C. 2007. Day length affects the dynamics of leaf expansion and cellular development in *Arabidopsis thaliana* partially through floral transition timing. *Ann Bot*. 99:703–711.

Debieu M, Tang C, Stich B, Sikosek T, Effgen S, Josephs E, Schmitt J, Nordborg M, Koornneef M, de Meaux J. 2013. Co-variation between seed dormancy, growth rate and flowering time changes with latitude in *Arabidopsis thaliana*. *PLoS One*. 8:e61075.

de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE. 2014. Genome scan methods against more complex models: when and how much should we trust them? *Mol Ecol*. 23:2006–2019.

Ding L, Kim SY, Michaels SD. 2013. FLOWERING LOCUS C EXPRESSOR family proteins regulate FLOWERING LOCUS C expression in both winter-annual and rapid-cycling *Arabidopsis*. *Plant Physiol*. 163:243–252.

Dunnett CW. 1955. A multiple comparison procedure for comparing several treatments with a control. *J Am Stat Assoc*. 50:1096–1121.

Ellegren H. 2014. Genome sequencing and population genomics in non-model organisms. *Trends Ecol Evol*. 29:51–63.

François O, Martins H, Caye K, Schoville SD. 2016. Controlling false discoveries in genome scans for selection. *Mol Ecol*. 25:454–469.

Geraldo N, Bäurle I, Kidou S, Hu X, Dean C. 2009. FRIGIDA delays flowering in *Arabidopsis* via a cotranscriptional mechanism involving direct interaction with the nuclear cap-binding complex. *Plant Physiol*. 150:1611–1618.

Gould BA, Stinchcombe JR. 2017. Population genomic scans suggest novel genes underlie convergent flowering time evolution in the introduced range of *Arabidopsis thaliana*. *Mol Ecol*. 26:92–106.

Gunther T, Coop G. 2013. Robust identification of local adaptation from allele frequencies. *Genetics*. 195:205–220.

Henderson IR, Dean C. 2004. Control of *Arabidopsis* flowering: the chill before the bloom. *Development*. 131:3829–3838.

Hohenlohe PA, Hand BK, Andrews KR, Luikart G. 2018. Population genomics provides key insights into ecology and evolution. In: Rajora OP, editor. *Population genomics: concepts, approaches, and applications*. Basel, Switzerland: Springer Nature Switzerland AG. p. 1–28.

Hohenlohe PA, Phillips PC, Cresko WA. 2010. Using population genomics to detect selection in natural populations: key concepts and methodological considerations. *Int J Plant Sci*. 171:1059–1071.

Joly-Lopez Z, Hoen DR, Blanchette M, Bureau TE. 2016. Phylogenetic and genomic analyses resolve the origin of important plant genes derived from transposable elements. *Mol Biol Evol*. 33:1937–1956.

Kim DH, Doyle MR, Sung S, Amasino RM. 2009. Vernalization: winter and the timing of flowering in plants. *Annu Rev Cell Dev Biol*. 25:277–299.

Kim HJ, Hyun Y, Park JY, Park MJ, Park MK, Kim MD, Kim HJ, Lee MH, Moon J, Lee I, *et al*. 2004. A genetic link between cold responses and flowering time through FVE in *Arabidopsis thaliana*. *Nat Genet*. 36:167–171.

Kingsley EP, Manceau M, Wiley CD, Hoekstra HE. 2009. Melanism in *Peromyscus* is caused by independent mutations in *Agouti*. *PLoS One*. 4:e6435.

Kleinboelting N, Huep G, Kloetgen A, Viehoever P, Weisshaar B. 2012. GABI-Kat SimpleSearch: new features of the *Arabidopsis thaliana* T-DNA mutant database. *Nucleic Acids Res*. 40:D1211–D1215.

Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu Rev Plant Biol*. 55:141–172.

Krysan PJ, Young JC, Sussman MR. 1999. T-DNA as an insertional mutagen in *Arabidopsis*. *Plant Cell*. 11:2283–2290.

Li YF, Costello JC, Holloway AK, Hahn MW. 2008. "Reverse ecology" and the power of population genomics. *Evolution*. 62:2984–2994.

Lotterhos KE, Whitlock MC. 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol Ecol*. 24:1031–1046.

Mackay TF, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet*. 10:565–577.

Martienssen RA. 1998. Functional genomics: probing plant gene function and expression with transposons. *Proc Natl Acad Sci USA*. 95:2021–2026.

Matus JT, Ferrier T, Riechmann JL. 2014. Identification of *Arabidopsis* knockout lines for genes of interest. *Methods Mol Biol*. 1110:347–362.

Michaels SD. 2009. Flowering time regulation produces much fruit. *Curr Opin Plant Biol*. 12:75–80.

Michaels SD, Amasino RM. 1999. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*. 11:949–956.

Moyle LC, Graham EB. 2005. Genetics of hybrid incompatibility between *Lycopersicon esculentum* and *L. hirsutum*. *Genetics*. 169:355–373.

Munguía-Rosas MA, Ollerton J, Parra-Tabla V, De-Nova JA. 2011. Meta-analysis of phenotypic selection on flowering phenology suggests that early flowering plants are favoured. *Ecol Lett*. 14:511–521.

O'Malley RC, Barragan CC, Ecker JR. 2015. A user's guide to the *Arabidopsis* T-DNA insertion mutant collections. *Methods Mol Biol*. 1284:323–342.

O'Malley RC, Ecker JR. 2010. Linking genotype to phenotype using the *Arabidopsis* unimutant collection. *Plant J*. 61:928–940.

Orr HA. 2009. Fitness and its role in evolutionary genetics. *Nat Rev Genet*. 10:531–539.

Østergaard L, Yanofsky MF. 2004. Establishing gene function by mutagenesis in *Arabidopsis thaliana*. *Plant J*. 39:682–696.

Pavey SA, Bernatchez L, Aubin-Horth N, Landry CR. 2012. What is needed for next-generation ecological and evolutionary genomics? *Trends Ecol Evol*. 27:673–678.

Roles AJ, Rutter MT, Dworkin I, Fenster CB, Conner JK. 2016. Field measurements of genotype by environment interaction for fitness caused by spontaneous mutations in *Arabidopsis thaliana*. *Evolution*. 70:1039–1050.

Rutter MT, Roles A, Conner JK, Shaw RG, Shaw FH, Schneeberger K, Ossowski S, Weigel D, Fenster CB. 2012. Fitness of *Arabidopsis thaliana* mutation accumulation lines whose spontaneous mutations are known. *Evolution*. 66:2335–2339.

Rutter MT, Shaw FH, Fenster CB. 2010. Spontaneous mutation parameters for *Arabidopsis thaliana* measured in the wild. *Evolution*. 64:1825–1835.

Rutter MT, Wieckowski YM, Murren CJ, Strand AE. 2017. Fitness effects of mutation: testing genetic redundancy in *Arabidopsis thaliana*. *J Evol Biol*. 30:1124–1135.

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

Samis KE, Murren CJ, Bossdorf O, Donohue K, Fenster CB, Malmberg RL, Purugganan MD, Stinchcombe JR. 2012. Longitudinal trends in climate drive flowering time clines in North American *Arabidopsis thaliana*. *Ecol Evol*. 2:1162–1180.

Sasaki E, Frommlet F, Nordborg M. 2018. GWAS with heterogeneous data: estimating the fraction of phenotypic variation mediated by gene expression data. *G3 (Bethesda)*. 8:3059–3068.

Schoville SD, Bonin A, François O, Lobreaux S, Melodelima C, Manel S. 2012. Adaptive genetic variation on the landscape: methods and cases. *Annu Rev Ecol Evol Syst*. 43:23–43.

Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, *et al*. 2002. A high-throughput *Arabidopsis* reverse genetics system. *Plant Cell*. 14:2985–2994.

Shindo C, Aranzana MJ, Lister C, Baxter C, Nicholls C, Nordborg M, Dean C. 2005. Role of FRIGIDA and FLOWERING LOCUS C in determining variation in flowering time of *Arabidopsis*. *Plant Physiol*. 138:1163–1173.

Skarnes WC, Rosen B, West AP, Koutsourakis M, Bushell W, Iyer V, Mujica AO, Thomas M, Harrow J, Cox T, *et al*. 2011. A conditional knockout resource for the genome-wide study of mouse gene function. *Nature*. 474:337–342.

Srikanth A, Schmid M. 2011. Regulation of flowering time: all roads lead to Rome. *Cell Mol Life Sci*. 68:2013–2037.

Stinchcombe JR, Hoekstra HE. 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity (Edinb)*. 100:158–170.

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

Talbot B, Chen TW, Zimmerman S, Joost S, Eckert AJ, Crow TM, Semizer-Cuming D, Seshadri C, Manel S. 2017. Combining genotype, phenotype, and environment to infer potential candidate genes. *J Hered*. 108:207–216.

Turner TL. 2014. Fine-mapping natural alleles: quantitative complementation to the rescue. *Mol Ecol*. 23:2377–2382.

Valentine ME, Wolyniak MJ, Rutter MT. 2012. Extensive phenotypic variation among allelic T-DNA inserts in *Arabidopsis thaliana*. *PLoS One*. 7:e44981.

Weigel D, Nordborg M. 2005. Natural variation in *Arabidopsis*. How do we find the causal genes? *Plant Physiol*. 138:567–568.

Weigel D, Nordborg M. 2015. Population genomics for understanding adaptation in wild plant species. *Annu Rev Genet*. 49:315–338.

Woody ST, Austin-Phillips S, Amasino RM, Krysan PJ. 2007. The WiscDsLox T-DNA collection: an *Arabidopsis* community resource generated by using an improved high-throughput T-DNA sequencing pipeline. *J Plant Res*. 120:157–165.