E-Article

Fitness Effects Associated with the Major Flowering Time Gene *FRIGIDA* in *Arabidopsis thaliana* in the Field

Tonia M. Korves,[1,*] Karl J. Schmid,[2,†] Ana L. Caicedo,[3,‡] Charlotte Mays,[3,§] John R. Stinchcombe,[4,‖] Michael D. Purugganan,[3,#] and Johanna Schmitt[1,]**

1. Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island 02912;
2. Department of Genetics and Evolution, Max-Planck Institute of Chemical Ecology, D-07745 Jena, Germany;
3. Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695;
4. Department of Botany, University of Toronto, Toronto, Ontario M5S 3B2, Canada

Submitted February 6, 2006; Accepted September 25, 2006; Electronically published March 7, 2007

Online enhancement: zip file.

ABSTRACT: To date, the effect of natural selection on candidate genes underlying complex traits has rarely been studied experimentally, especially under ecologically realistic conditions. Here we report that the effect of selection on the flowering time gene *FRIGIDA* (*FRI*) reverses depending on the season of germination and allelic variation at the interacting gene *FLOWERING LOCUS C* (*FLC*). In field studies of 136 European accessions of *Arabidopsis thaliana*, accessions with putatively functional *FRI* alleles had higher winter survival in one *FLC* background in a fall-germinating cohort, but accessions with deletion null *FRI* alleles had greater seed production in the other *FLC* background in a spring-germinating cohort. Consistent with *FRI*'s role in flowering, selection analyses suggest that the difference in winter survival can be attributed to time to bolting. However, in the spring cohort, the fitness difference was associated with rosette size. Our analyses also reveal that controlling for population structure with estimates of inferred ancestry and a geographical restriction was essential for detecting fitness associations. Overall, our results suggest that the combined effects of seasonally varying selection and epistasis could explain the maintenance of variation at *FRI* and, more generally, may be important in the evolution of genes underlying complex traits.

Keywords: epistasis, *FLOWERING LOCUS C* (*FLC*), population structure, vernalization, candidate gene association, heterogeneous selection.

* E-mail: tonia_korves@brown.edu.

† Present address: Genebank Department, Leibniz Institute of Plant Genetics and Cultivated Plant Research, D-06466 Gatersleben, Germany; e-mail: karl@minzer-schmid.de.

‡ Present address: Biology Department, University of Massachusetts, Amherst, Massachusetts 01003; e-mail: caicedo@bio.umass.edu.

§ E-mail: flower@indythinker.com.

‖ E-mail: stinchcombe@botany.utoronto.ca.

Present address: Department of Biology and Center for Comparative Functional Genomics, New York University, New York, New York 10003; e-mail: mp132@nyu.edu.

** E-mail: johanna_schmitt@brown.edu.

Am. Nat. 2007. Vol. 169, pp. 000–000. 0003-0147/2007/16905-41615$15.00.

How natural selection on complex traits affects allelic variation at the genes underlying these traits has long been a central question in evolutionary biology (Fisher 1930; Wright 1931; Gillespie and Turelli 1989; Whitlock et al. 1995; Weinreich et al. 2005). Only recently has it become possible to evaluate mechanisms of selection affecting known genes and genetic pathways underlying complex developmental traits, as the genes and pathways that contribute to natural variation in such traits have been identified (Hanson et al. 1996; Johanson et al. 2000; Long et al. 2000; Glazier et al. 2002). Patterns of nucleotide polymorphism at several developmental genes strongly suggest that natural selection has acted to maintain variation at these loci (Olsen et al. 2002, 2004; Wright and Gaut 2005; Toomajian et al. 2006; Voight et al. 2006). So far, few studies have examined the mechanisms of natural selection that affect candidate polymorphisms under ecologically realistic conditions (Watt 1977; Eanes 1999; Tian et al. 2003; Hoekstra et al. 2004), especially for developmental genes underlying complex traits.

Evolutionary theory suggests several possible mechanisms for the maintenance of genetic polymorphism, both within populations and within species. Epistatic selection, in which the fitness effects of alleles at one locus depend on which alleles are present at other loci, may maintain polymorphism within populations (Gimelfarb 1989), contribute to variation between populations (Wright 1931; Wade and Goodnight 1998), and constrain evolutionary

trajectories (Weinreich et al. 2005). Several recent studies of quantitative trait loci (QTLs) suggest that epistasis may be an important source of fitness variation (Malmberg et al. 2005) and may contribute to patterns of nucleotide polymorphism indicative of balancing selection (Weinig et al. 2003; Kroymann and Mitchell-Olds 2005). Heterogeneous selection favoring different alleles in different environments may also maintain genetic variation (Levene 1953; Gillespie and Turelli 1989; Ellner and Hairston 1994) both within populations, through temporal variation in selection, and among populations, through local adaptation to different selective pressures in different sites.

Flowering time in the annual plant *Arabidopsis thaliana* is an ideal system for studying mechanisms of selection affecting known genes because many of the genes involved and their interactions are known (Simpson and Dean 2002; Boss et al. 2004), polymorphic genes that contribute to flowering time variation have been identified (Johanson et al. 2000; Gazzani et al. 2003; Michaels et al. 2003; Caicedo et al. 2004; Olsen et al. 2004; Lempe et al. 2005; Shindo et al. 2005; Werner et al. 2005*b*), and fitness estimates can be obtained in field conditions (Weinig et al. 2003). Natural populations of *A. thaliana* experience a wide range of climatic conditions across the species' geographic range that are likely to exert very different selective pressures on seasonal timing. In many locations, *A. thaliana* plants exhibit winter annual behavior, in which seedlings germinate in the fall, overwinter as small vegetative rosettes of leaves, and flower and set seed in spring. Depending on the particular conditions, plants can also behave as spring or summer annuals, germinating in the spring or summer and flowering in spring, summer, or fall. In populations with several generations per year, different seasonal cohorts may experience very different selection on complex trait variation (Donohue 2002; Weinig et al. 2003; Donohue et al. 2005).

A major contributor to flowering time variation in *A. thaliana* is the gene *FRIGIDA* (*FRI*; Napp-Zinn 1987; Johanson et al. 2000). Functional *FRI* alleles cause plants to delay flowering if they have not experienced a period of cold, a process known as vernalization (Napp-Zinn 1987). Nonfunctional *FRI* alleles cause plants to flower rapidly in the absence of vernalization (Johanson et al. 2000). Nonfunctional *FRI* alleles have arisen multiple times, and two null deletion alleles occur with high frequencies (Johanson et al. 2000; Le Corre et al. 2002; Gazzani et al. 2003; Hagenblad et al. 2004; Lempe et al. 2005; Shindo et al. 2005; Werner et al. 2005*a*). Analyses of DNA sequence variation at the *FRI* locus suggest a rapid, recent increase in the frequency of null alleles, consistent with positive selection (Johanson et al. 2000; Le Corre et al. 2002; Le Corre 2005; Toomajian et al. 2006). In addition, in a set of French populations, *FRI* functional variation was more differentiated between populations than expected from markers, suggesting that there is local selection on *FRI* functionality (Le Corre 2005).

There are several hypotheses about the selective agents affecting *FRI* variation. *FRI* functional alleles are believed to be favored for winter annual behavior because they delay flowering until the appropriate time in spring (Simpson and Dean 2002), and null alleles are thought to be favored for summer annual behavior (Johanson et al. 2000; Pigliucci and Marlow 2001). It has also been proposed that selection on *FRI* varies with *FLOWERING LOCUS C* (*FLC*) genotype (Caicedo et al. 2004). *FRI* upregulates *FLC*, a MADS-box transcriptional activator that inhibits flowering and is downregulated by vernalization (Michaels and Amasino 1999). The gene *FLC* has two major haplogroups that differ in intronic regions and in a radical amino acid change in an alternatively spliced transcript induced at high levels by vernalization (Caicedo et al. 2004). Though on different chromosomes, some combinations of *FRI* and *FLC* alleles are under- or overrepresented among European accessions, suggesting epistatic selection (Caicedo et al. 2004).

Here we evaluate selection on *FRI* in the field by examining associations between *FRI FLC* genotypes and fitness traits in 136 *A. thaliana* accessions. Because, in association studies, cryptic population structure can lead to spurious candidate gene associations (Cardon and Palmer 2003), we employed strict controls for population structure. We report that fitness effects associated with *FRI* depend on the *FLC* genetic background as well as on the seasonal environment.

Material and Methods

Plant Material

Arabidopsis thaliana, commonly known as thale cress or mouse-ear cress, is an annual weed native to Eurasia and now widely found in North America. Although *A. thaliana* is highly self-fertilizing, genetic variation is found both between and within *A. thaliana* populations (Le Corre 2005; Stenoien et al. 2005; Bakker et al. 2006). When *A. thaliana* plants are collected from the wild, the selfed seed can be maintained as a line known as an "accession."

We chose 360 *A. thaliana* accessions based on what was available at the time from the *Arabidopsis* Biological Resource Center (Ohio State University) and used these in a field experiment. In our analyses, in order to control for population structure, we included only the accessions for which we had single-nucleotide polymorphism (SNP) genotype data (Schmid et al. 2006). In order to consider each accession as an independent sample, we included only one accession per unique SNP genotype per collection

location within Europe (west of 40°E longitude); 30 collection locations gave rise to more than one genotype and were represented by more than one accession. We also excluded accessions that had rare insertion/deletion *FRI* alleles (i.e., alleles unlike the *FRI* alleles found in the accessions Sf-2, Columbia, and Ler) that we were able to detect (Stinchcombe et al. 2004). These criteria resulted in a set of 169 accessions in our analyses. The accession data are available in a zip archive, in both an Excel file and tab-delimited ASCII. The set used was further reduced to 136 accessions in analyses where an additional control for population structure was implemented (see "Controls for Population Ancestry").

Field Experiment

We planted the accessions in the fall of 2002 and spring of 2003 in a plowed field at Brown University's Haffenreffer Reserve in Bristol, Rhode Island. The field site is in a region where both fall-germinating winter annual and spring-germinating summer annual cohorts occur in wild *A. thaliana*. Plowed fields are a typical habitat for this ruderal species; wild *A. thaliana* populations often occur in agricultural fields (Le Corre 2005; T. Korves, personal observation). This is the first publication of results from this field experiment; previous studies have used the same field site (Weinig et al. 2002; Donohue et al. 2005) or a subset of the same accessions at another Rhode Island field site (Stinchcombe et al. 2004). The field experiment reported here was conducted during a much colder winter than these previous field experiments (based on monthly averages available at http://www.erh.noaa.gov/box/dailystns.shtml), which may be why we observed higher winter mortality (see "Results").

Seeds were planted in Metromix 360 soil in 96 cell flats and cold stratified in the dark at 4°C for 4 days for the fall planting and for 2 weeks for the spring planting. Seedlings were germinated and grown in a greenhouse for 2 weeks and transferred to cold frames for several days before planting in the field. Seedlings were transplanted with their soil plugs into the field on October 28–31 and April 1–3.

For each the and spring cohort, we planted 10 blocks, with one replicate of each accession per block. Each block contained 360 accessions (of which only 169 are used in the analyses presented in this article) in an 8 × 45 plant grid with 10-cm spacing between plants. This spacing is within the range of density observed in wild populations (T. Korves, personal observation). Accessions were randomly assigned to positions within each block. Blocks for the fall and spring generations were interspersed within the field. Only those plants that survived transplanting in the field were included in analyses.

Because *A. thaliana* is primarily self-fertilizing, we assessed fitness by estimating total seed mass, calculated as the product of the number of fruits and an estimate of seed mass per fruit (see app. A for more details). In the fall planting, there was 66% mortality over the winter, and consequently total seed mass was not unimodal. Therefore, we evaluated fall fitness in two components: winter survival and total seed mass per winter survivor. To assess winter survival, we scored whether plants were alive or dead after snowmelt on March 25. Because prewinter rosette size may affect winter survival, we measured rosette diameter on December 13 and 14 in the fall cohort. Because rosette size at the time of bolting has been associated with seed production (Griffith et al. 2004), we measured rosette diameter at the time of bolting in the spring cohort. Bolting was assessed every few days in the spring and every 2–3 weeks during the winter when snow cover permitted.

Genotyping

FRI genotyping methods are described by Stinchcombe et al. (2004). *FRI* alleles were considered putatively functional if they did not have any of the three deletions tested for or unusual polymerase chain reaction (PCR) products. Putatively functional *FRI* alleles hereafter are referred to as *FRI*, *FRI* alleles containing a 376-bp deletion, as in the Landsberg *erecta* accession, are referred to as FRI^{delLer}, and *FRI* alleles containing a 16-bp deletion, as in the Columbia accession, are referred to as FRI^{delCol} (Johanson et al. 2000; Stinchcombe et al. 2004). It is possible that some accessions classified as having putatively functional *FRI* alleles in our study have rare *FRI* null alleles that we were unable to detect with our PCR markers. However, the accessions we identified as having putatively *FRI* functional alleles did not include any accessions that were identified as having rare *FRI* null alleles in three recent studies (Lempe et al. 2005; Shindo et al. 2005; Werner et al. 2005*a*). For *FLC*, accessions were genotyped for the two major haplogroups, FLC^{A} and FLC^{B}, and for common insertions in intron 1, as described by Caicedo et al. (2004). *FRI* and *FLC* genotypes are provided in the zip archive. To account for population structure, we used SNP data for 115 markers; the collection of these data is described by Schmid et al. (2006).

Controls for Population Ancestry

Arabidopsis thaliana displays population structure (Sharbel et al. 2000; Nordborg et al. 2005; Schmid et al. 2006). Because cryptic population structure can lead to spurious candidate gene associations (Cardon and Palmer 2003), we controlled for population structure in two ways: by estimating population ancestry and by using a geographically restricted sample. The inferred-ancestry estimates

were used to remove some of the variation between different genetic backgrounds that was not due to the candidate genes, much like blocking factors. This should have the effect of both reducing bias due to population structure and reducing noise variation, thereby increasing the power to detect associations. We also used a geographical restriction because, beyond population structure captured by randomly chosen marker loci, there may be loci under selection for adaptation to climate that could result in spurious associations or obscure real associations. This was of particular concern for our study because our candidate allelic variation was not evenly distributed geographically across all of Europe and because geographic and climatic variables are associated with fitness traits (T. Korves, unpublished manuscript). Thus, eliminating accessions from geographical regions where only a subset of the candidate allelic variation is present might help to reduce biases due to other genes under selection.

To estimate the population ancestry of each accession, we used SNP data (Schmid et al. 2006) and the program *structure* 2.0 (Pritchard et al. 2000*a*, 2000*b*). To determine the most appropriate number of populations, *K*, for our set of 169 accessions, we ran the model five times for each *K*, $K = 1$–7, calculated likelihoods of the data, given *K*, and checked for consistency across runs. We chose $K = 6$, based on the highest estimated log likelihood of the data, given *K* (Pritchard et al. 2000*b*; table A1). We used the inferred-ancestry estimates from one $K = 6$ *structure* run as covariates in association analyses. Because of extensive admixture in our sample, the inferred-ancestry estimates are estimates of the proportion of each accession's genome that came from each of the inferred ancestral populations. The inferred-ancestry estimates are given in the zip archive. Further details about how we estimated ancestry and figures displaying the inferred-ancestry values using different numbers of populations, created with the program *Distruct* (Pritchard et al. 2000*b*), are provided in appendix A.

We geographically restricted the set of accessions to those from the northwestern European region (between 44° and 54°N latitude and west of 22°E longitude) where *FRI* null alleles are common and where we had the greatest sampling density (see fig. A1 for a map). Within the geographically restricted region, both *FLC* haplogroups are present, and neither shows a latitudinal cline in frequency in accessions with functional *FRI* alleles (one-way ANOVA with latitude and *FLC* haplogroup: $F = 0.05$, $\mathrm{df} = 26$, $P = .83$), as they do across a greater range of latitudes in Europe (Caicedo et al. 2004). The geographical restriction resulted in a set of 136 accessions for genotype-trait association analyses. We did not choose more-restrictive latitudes and longitudes in order to maintain a sufficient sample size for analysis.

We examined the effects of these controls for population structure on our results by performing analyses with and without each of these controls. In addition, because not all of the ancestries present among deletion null *FRI* FLC^{B} accessions were well represented among *FRI* FLC^{B} accessions (for which there were only eight accessions), we examined the effect of restricting the data set for better ancestry matching and found that this did not affect our association results (see app. B). We also examined the effect of restricting the sample to central European accessions and found that this did not qualitatively affect the association results (see app. B).

Statistical Analyses

We compared accessions with functional *FRI* alleles to those with the two common *FRI* deletion null alleles, FRI^{delLer} and FRI^{delCol}, and examined associations across and within *FLC* haplotype backgrounds. Associations were evaluated using the ANOVA model: accession trait mean = *FRI* functionality × *FLC* + *FRI* functionality + *FLC* haplogroup + population 1 ancestry + population 2 ancestry + … + population 5 ancestry. Ancestry from the sixth population is collinear with the sum of the other ancestries and thus was not included in the analyses. *FRI* functionality had two categories, *FRI* functional and *FRI* null, in which the two types of deletion alleles were pooled together; this pooling was done in order to evaluate epistasis with *FLC*. To test the hypothesis that *FRI* functional and *FRI* null alleles differed, we replaced *FRI* functionality with *FRI* allele (which included the three classes: *FRI* functional, FRI^{delLer}, and FRI^{delCol}) and compared *FRI* functional alleles with the FRI^{delLer} and FRI^{delCol} null alleles in means contrasts. To further investigate the probability of observing genetic locus–fitness trait associations like those we found, we performed similar analyses using the SNPs. We compared the *F* statistics from means contrasts between SNP alleles within an *FLC* background with the *F* statistic from the means contrast of *FRI* functional alleles versus the two null allele classes. Then we calculated the percentage of SNPs that yielded higher test statistics than *FRI* (e.g., Thornsberry et al. 2001). For this, we used only SNPs with allele frequencies greater than 5% and sufficient variation for interactions. To determine the percentage of variation genotype explained, we calculated η^2 values (sum of squares for genotype effects divided by total sum of squares). The percentage of variation explained by ancestry covariates was calculated similarly.

To examine whether *FRI FLC* genotype may have affected fitness through effects on time to bolting, we performed multivariate genotypic selection analyses, using accession trait means (Lande and Arnold 1983; Rausher 1992; Stinchcombe et al. 2002). In the genotypic selection

analysis for winter survival in the fall cohort, we included time to bolting, because delayed bolting increases winter survival in *A. thaliana*'s relative, *Brassica* (OMAF 2002), and prewinter rosette diameter, a trait known to be positively associated with survival and seed production in *A. thaliana* (Griffith et al. 2004). In the spring cohort, we included time to bolting, because of hypotheses about the effects of bolting time on spring fitness (Johanson et al. 2000; Pigliucci and Marlow 2001), and rosette diameter at bolting, because rosette diameter is associated with seed production in *A. thaliana* (Griffith et al. 2004). We also analyzed *FRI FLC* genotype associations for time to bolting and rosette diameter, using the same model as for the fitness traits. For the analysis of rosette diameter in the spring cohort, we included time to bolting as a covariate to factor out differences in the timing of the measurement of rosette diameter; rosette diameter was measured at the time of bolting, and time to bolting and rosette diameter at bolting were highly, positively correlated (see "Results"; fig. 2*B*).

We used accession means in our analyses instead of raw data in order to include the inferred-ancestry covariates, which are properties of lines and not of individual plants and therefore could not be used with raw data. We used accession least squares means for total seed mass and rosette diameter, which were calculated with ANOVAs that included block. Because survival is a binary trait, we used a logit model that included accession and block to calculate predicted winter survival for the accessions. For time to bolting, we used medians because bolting dates were not normally distributed within genotypes. For time to bolting in the fall cohort, because of high winter mortality, we used only accessions for which we were able to observe bolting for at least three plants. To meet the normality requirements of ANOVA, least squares means for total seed mass for fall-cohort winter survivors were log transformed, and median bolting times in the fall cohort were transformed with a box-cox transformation. The sample size was 136 accessions, except where we note otherwise in the results. Statistical analyses were done using Statistica 6.0 and Intercooled Stata 8. Accession means are provided in the zip archive.

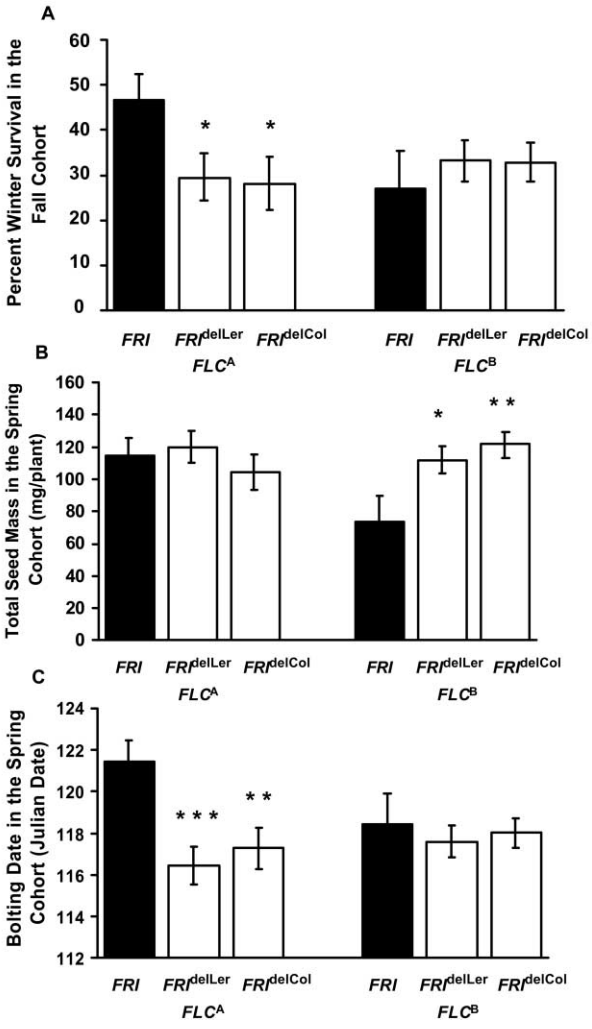

Figure 1: *A*, Predicted means (±1 SE) for percent winter survival in the fall cohort. *FRI* (*FRIGIDA*) functionality × *FLC* (*FLOWERING LOCUS C*): $F = 4.56$, df $= 1, 127$, $P = .035$. *B*, Least squares means for total seed mass per plant in the spring cohort. *FRI* functionality × *FLC*: $F = 4.74$, df $= 1, 127$, $P = .031$. *C*, Least squares means for time to bolting in the spring cohort. *FRI* functionality × *FLC*: $F = 4.74$, df $= 1, 127$, $P = .031$. Asterisks indicate significant differences between *FRI* deletion alleles (FRI^{delLer} and FRI^{delCol}) and *FRI* functional alleles (*FRI*) within an *FLC* haplogroup based on means contrasts: one asterisk for $P < .05$, two asterisks for $P < .01$, and three asterisks for $P < .0001$.

Results

Associations between FRI *Allelic Variation and Fitness Traits*

In the fall cohort, *FRI* variation was associated with winter survival, but in only one *FLC* background (fig. 1*A*). Accessions with *FRI* functional alleles had, on average, 1.6 times higher winter survival than accessions with null alleles in the FLC^A background (means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 6.79$, df $= 1, 125$, $P = .010$), but there was no difference in the FLC^B background ($F = 0.44$, df $= 1, 125$, $P = .51$). *FRI*, *FLC*, and their interaction together accounted for 5.4% of the variation among accessions in winter survival, and within the FLC^A background, *FRI* explained 13.5% of the variation. Because insertions in the *FLC* intron 1 are known to cause weak alleles of *FLC* (Gazzani et al. 2003; Michaels et al. 2003), we repeated the analysis excluding all accessions that have

insertions in *FLC* intron 1 and found that these insertions are not responsible for the *FRI* association ($N = 124$; *FRI* functionality × *FLC*: $F = 4.08$, df $= 1, 115$, $P = .046$; means contrast *FRI* vs. FRI^{delLer} and FRI^{delCol} in FLC^{A}: $F = 4.73$, df $= 1, 113$, $P = .032$). We detected no differences between *FRI* functional and *FRI* null allele accessions in total seed mass among winter survivors ($N = 123$ accessions; *FRI* functionality × *FLC*: $F = 0.002$, df $= 1, 114$, $P = .96$; means contrast *FRI* vs. FRI^{delLer} and FRI^{delCol} in FLC^{A}: $F = 0.003$, df $= 1, 112$, $P = .96$; in FLC^{B}: $F = 0.09$, $P = .76$).

In the spring cohort, *FRI* functionality was also associated with fitness; this association also depended on *FLC* background but in the opposite direction (fig. 1*B*). *FRI* functional alleles were associated with, on average, 38% lower total seed mass than the *FRI* null alleles in the FLC^{B} background (means contrast: $F = 6.38$, df $= 1, 125$, $P = .013$), and there was no difference in the FLC^{A} background ($F = 0.03$, df $= 1, 125$, $P = .86$). *FRI*, *FLC*, and their interaction together accounted for 8.0% of the variation among accessions in seed production, and within the FLC^{B} background, *FRI* explained 9.2% of the variation. The difference in total seed mass in FLC^{B} was due to nonfunctional alleles being associated with both a greater number of fruits (means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles in FLC^{B}: $F = 4.46$, df $= 1, 125$, $P = .037$) and greater seed mass per fruit ($F = 7.77$, df $= 1, 125$, $P = .006$).

Associations like those observed for *FRI* within *FLC* backgrounds were rare at other loci, suggesting that these genetic associations were unlikely to be observed by chance. In the fall cohort in the FLC^{A} background, only three SNP markers (3.8%) yielded a higher *F* statistic for winter survival variation than *FRI*. In the spring cohort in the FLC^{B} background, only one SNP (1.3%) yielded a higher *F* statistic for seed mass variation than *FRI*. In the other *FLC* background in each cohort, SNP allele associations of greater significance than those of *FRI* were also rare (1 of 78 for the fall and 0 of 78 for the spring). However, these results should be treated with caution because the number of SNPs used in these analyses is not especially large, and there are differences in statistical power for markers with different allele frequencies.

Association between FRI *Allelic Variation and Time to Bolting in the Spring Cohort*

In the spring cohort, *FRI* functional alleles were associated with delayed bolting in the FLC^{A} background but not in the FLC^{B} background (fig. 1*C*). *FRI*, *FLC*, and their interaction together accounted for 9.2% of the variation among accessions in time to bolting, and within the FLC^{A} background, *FRI* explained 14.4% of the variation. None of 78 SNP markers yielded a higher *F* statistic for means comparisons within FLC^{A} or FLC^{B} than *FRI*, suggesting that this result was unlikely to be observed by chance. We performed a more limited analysis for time to bolting in the fall cohort because our data were incomplete as a result of snow cover and high mortality (see next subsection).

Selection Mechanisms behind Genotype-Fitness Associations

Because *FRI* and *FLC* affect bolting (Michaels and Amasino 1999; Johanson et al. 2000), we examined whether the *FRI FLC* associations with the fitness measures can be attributed to differences in time to bolting. Our bolting data for the fall cohort were incomplete because the plants were under snow cover for much of the winter (making bolting unobservable), most plants died before snow melt, and bolting before the snow cover was rare. Consequently, our measurements of time to bolting may be biased because of missing data, and associations with time to bolting in our experiment must be treated with caution. However, time to bolting in our experiment did correlate highly with bolting data from a previous experiment (Stinchcombe et al. 2004), planted in the fall and conducted in the field in Rhode Island, that experienced substantially lower winter mortality ($N = 19$ accessions overlapping between the experiments; accession means from Stinchcombe et al. 2004; $R = 0.80$, $P < .001$).

Our results indicate that delayed bolting contributed to increased winter survival. Among the 76 accessions with at least three plants surviving until bolting, when differences in prewinter size were accounted for, accessions that bolted later had higher winter survival (fig. 2*A*). Consistent with this, accessions with at least one replicate that bolted before snow cover had lower winter survival (one-way ANOVA, $F = 8.56$, df $= 1, 134$, $P = .004$).

FRI functional alleles were associated with delayed bolting (fig. 2*B*). Accessions with *FRI* functional alleles bolted later than those with null alleles in both the FLC^{A} (means contrast: $F = 9.25$, df $= 1, 65$, $P = .0034$) and FLC^{B} backgrounds (means contrast: $F = 4.21$, df $= 1, 65$, $P = .044$). None of the SNPs (0 of 78) had a higher *F* statistic for an association with time to bolting than *FRI*. There was no main effect of *FLC* or *FRI* × *FLC* interaction (*FLC*: $F = 0.31$, df $= 1, 67$, $P = .58$; *FRI* functionality × *FLC*: $F = 0.02$, df $= 1, 67$, $P = .87$). *FRI FLC* genotype was not associated with prewinter rosette diameter (*FRI* functionality × *FLC*; $F = 2.88$, df $= 1, 127$, $P = .53$; *FRI* main effect: $F = 4.82$, df $= 1, 127$, $P = .42$; *FLC* main effect: $F = 3.39$, df $= 1, 127$, $P = .49$). Together, these results suggest that *FRI* delayed bolting and that this delay in bolting contributed to increased winter survival.

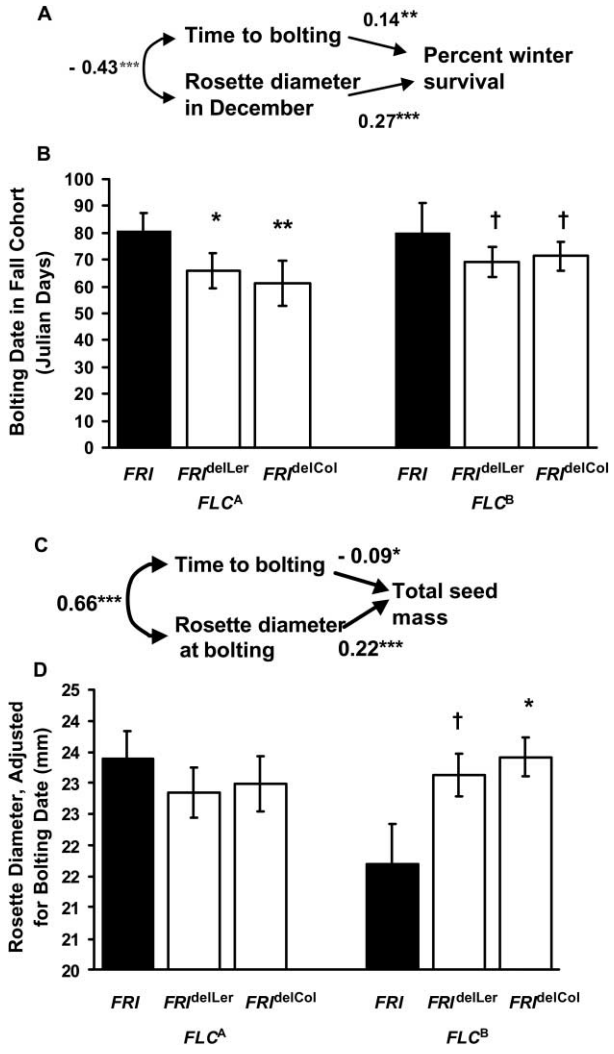

Figure 2: *A*, Standardized selection gradients (β) and correlation for the fall cohort ($N = 76$ accessions). *B*, Least squares means (± 1 SE) for time to bolting in the fall cohort. Main effect of *FRIGIDA* (*FRI*) functionality: $F = 12.93$, df $= 1, 67$, $P = .0006$. Bolting-time measurements in *A* and *B* are for lines with at least three individuals surviving until bolting. *C*, Standardized selection gradients (β) and correlation for the spring cohort ($N = 136$ accessions). *D*, Least squares means (± 1 SE) for rosette diameter at bolting in the spring cohort, adjusted for time to bolting. *FRI* functionality × *FLC* (*FLOWERING LOCUS C*): $F = 6.19$, df $= 1, 126$, $P = .014$. In *B* and *D*, markers indicate differences between *FRI* deletion alleles and *FRI* functional alleles within an *FLC* haplogroup based on means contrasts: cross for $P < .10$, one asterisk for $P < .05$, two asterisks for $P < .01$, and three asterisks for $P < .001$.

However, because the delay in bolting occurred in both *FLC* backgrounds, these results do not explain why higher winter survival was observed only in the FLC^A background and not in the FLC^B background.

In contrast, selection in the spring cohort favored earlier bolting and larger rosette diameter at bolting (fig. 2*C*). Because *FRI* genotype was not associated with a difference in time to bolting in the FLC^B background (fig. 1*C*), where the difference in seed production was observed, selection on time to bolting cannot explain the selection against functional *FRI* in the FLC^B background. Instead, the genotypic difference in seed production may be explained by variation in rosette diameter, a trait that had a stronger effect than time to bolting on seed production (fig. 2*C*). The association of *FRI* with rosette diameter varied by *FLC* class and parallels the results for seed production (fig. 2*D*). In the FLC^B background, *FRI* functional accessions had significantly smaller rosettes than *FRI* null allele accessions (means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles in FLC^B: $F = 5.25$, df $= 1, 124$, $P = .024$; 3.8% of SNP markers had a stronger association with rosette diameter than *FRI* within FLC^B). These results indicate that *FRI* FLC^B accessions had slower rosette growth and that this smaller rosette size may have led to lower seed production in the spring cohort.

Effects of Controls for Population Structure

There is controversy over the importance of controlling for population structure in association analyses (Cardon and Palmer 2003). Therefore, we examined the importance of the two controls for population structure that we used in our study: estimating inferred population ancestry and restricting the set of accessions to a geographical region. Both of these controls were necessary for detecting epistatic associations for the fitness traits and for time to bolting in the spring cohort (table 1, lines 1, 2, 9, and 10).

The effect of the geographical restriction for winter survival and for seed production in the spring cohort was due in large part (though not entirely) to the removal of southern accessions, those south of 44°N latitude (table 1, lines 3, 4). The reason is that the inclusion of southern accessions obscured associations present within the geographically restricted region. The southern accessions had lower winter survival and seed production in the spring cohort than accessions from north of 44°N latitude (one-way ANOVA; winter survival: $F = 20.28$, df $= 1, 167$, $P < .00001$; seed production: $F = 3.99$, df $= 1, 167$, $P = .047$) and were predominantly *FRI* FLC^A (16 of the 19 southern accessions were *FRI* FLC^A). In contrast, in the geographically restricted region, *FRI* FLC^A was associated with high winter survival and was not associated with low seed production in the spring. Because our sample of accessions from south of 44°N had little *FRI* and *FLC* variation, we cannot assess whether *FRI* or *FLC* is associated with the traits we measured in southern accessions.

Within the geographically restricted set of accessions, the inclusion of inferred-ancestry covariates enhanced the power to detect *FRI* × *FLC* associations despite the fact

Table 1: *FRI* functionality × *FLC* results with and without geographical restrictions and the inferred-ancestry covariates in the ANOVA model

Line	Geographical restriction	Ancestry covariates	N	df	Winter survival		Spring seed mass		Time to bolting in spring cohort	
					F	P	F	P	F	P
1	44°–54°N latitude, west of 22°E longitude	Yes	136	1, 127	4.56	.03	4.74	.03	4.74	.03
2	44°–54°N latitude, west of 22°E longitude	No	136	1, 132	2.78	.10	2.93	.09	2.41	.12
3	North of 44°N latitude	Yes	150	1, 141	2.18	.14	3.81	.05	.67	.42
4	North of 44°N latitude	No	150	1, 146	1.24	.27	2.12	.15	.15	.70
5	South of 54°N latitude	Yes	159	1, 150	1.66	.20	1.43	.23	1.74	.19
6	South of 54°N latitude	No	159	1, 155	.05	.83	.26	.26	.06	.80
7	West of 22°E longitude	Yes	158	1, 149	.81	.37	2.31	.13	.06	.81
8	West of 22°E longitude	No	158	1, 154	.00	.95	1.89	.17	.28	.60
9	None	Yes	169	1, 160	.35	.56	1.52	.22	.00	.99
10	None	No	169	1, 165	.10	.75	1.29	.26	.42	.52

that these covariates did not explain an especially large amount of trait variation. In the model with *FRI*, *FLC*, and *FRI* × *FLC* with the geographically restricted set of accessions, the ancestry covariates together accounted for just 2.8% of the variation in winter survival, 6.3% of the variation in seed production in the spring cohort, and 7.1% of the variation in time to bolting in the spring cohort. When no *FRI* and *FLC* genotype effects were included in the model, the ancestry covariates together did not explain a significant amount of the variation in either fitness trait or in time to bolting in the spring (adjusted R^2 values, all $P > .2$). However, for seed production in the spring cohort, greater ancestry from inferred ancestral population 5 (see "Methods for Estimating Population Ancestry" in app. A for more information on the inferred populations) was marginally associated with higher seed production ($R = 0.17$, $P = .043$), and greater ancestry from inferred ancestral population 4 was marginally associated with lower winter survival ($R = -0.17$, $P = .050$).

Inclusion of inferred ancestry from a single population, population 2, was sufficient for creating *FRI* × *FLC* associations with $P < .05$ for both fitness traits and for time to bolting in the spring cohort. In a model with *FRI*, *FLC*, and their interaction, population 2 ancestry explained 3.2% of time to bolting variation in the spring ($F = 4.71$, df $= 1, 131$, $P = .032$). However, population 2 ancestry was not significantly associated with winter survival or with seed production in the spring cohort (in models with *FRI*, *FLC*, and *FRI* × *FLC*: $F = 2.0$, df $= 1, 131$, $P = .16$ and $F = 1.4$, df $= 1, 131$, $P = .23$, respectively). Population 2 ancestry was significantly correlated with latitude ($R = 0.46$, $P < .0001$), suggesting that it may remove population structure effects on time to bolting associated with latitude. No other population ancestry was sufficient on its own for reducing *FRI* × *FLC* P values below 0.05 for any of the three traits.

Whether or not the geographical restriction and/or the inferred-ancestry covariates were included had no effect on the association we found for time to bolting in the fall cohort; *FRI* was always associated with longer time to bolting, and *FRI* × *FLC* and *FLC* haplogroup had no associations (results not shown). These results contrast with an earlier result from a previous experiment, also planted in the fall and conducted in the field in Rhode Island (Stinchcombe et al. 2004), in which *FRI FLC*^{A} was associated with early bolting (Caicedo et al. 2004). The previous study employed no geographical restrictions and analyzed a set of accessions that had no detectable population structure based on an amplified fragment length polymorphism (AFLP) data set (Olsen et al. 2004). We reanalyzed the data set used by Caicedo et al. (2004) with our geographical restrictions and our inferred-ancestry estimates based on the SNPs and found that these controls for population structure reversed the direction of the association of *FRI FLC*^{A} with time to bolting (table B1). With these controls for population structure, accessions with *FRI* bolted later than accessions with *fri* in the *FLC*^{A} background, as in our experiment. Further analyses indicate that a major factor contributing to the different results with and without our controls for population structure is the inclusion of accessions from Spain (see app. B).

We also examined whether population structure, as determined from the SNP data, could explain another result from a previous study (Stinchcombe et al. 2004) that used the same time-to-bolting data as in Caicedo et al. (2004). We found that the inclusion of our inferred-ancestry covariates could not account for a latitudinal cline specific to accessions with *FRI* functional alleles,

suggesting that this cline is not due to population structure (see app. B).

Discussion

Genomic studies of natural variation in model organisms are now making it possible to identify the genes underlying complex trait variation and to investigate mechanisms of natural selection that affect those genes and traits. In *Arabidopsis thaliana*, genome-wide scans have successfully identified such genes via QTL mapping in recombinant inbred lines (El-Assal et al. 2001; Kroymann and Mitchell-Olds 2005; Werner et al. 2005*b*) and confirmed them via association mapping in accessions (Aranzana et al. 2006). A complementary approach is to examine allelic variation at candidate genes of known function, identified via molecular genetic studies of mutant and transgenic plants. Several recent studies have detected nonneutral patterns of sequence polymorphism at such candidate genes, suggesting that they may have been affected by historical selection (Stahl et al. 1999; Le Corre et al. 2002; Olsen et al. 2002; Tian et al. 2002; Mauricio et al. 2003; Schmid et al. 2005; Toomajian et al. 2006). However, to understand the ecological mechanisms that affect candidate genes, it is necessary to measure selective forces in real time under field conditions. Our results demonstrate that allelic variation in the important candidate gene *FRI* is associated with fitness variation under field conditions and that which type of allele is favored depends on the seasonal environment and the genetic background. These results add to the growing evidence that epistatic selection may be an important mechanism for maintaining genetic variation in *A. thaliana* and other species (Routman and Cheverud 1997; Shook and Johnson 1999; Leips and Mackay 2000; Weinig et al. 2003; Peripato et al. 2004; Kroymann and Mitchell-Olds 2005; Malmberg et al. 2005).

Our study was conducted with European ecotypes within the introduced North American range of *A. thaliana*, in a region where spring and fall seasonal cohorts are commonly observed within the same populations. Our results predict that functional *FRI FLC*^{A} genotypes may have been favored and that *FRI FLC*^{B} genotypes may have been selected against during colonization of New England by European genotypes. Because our study was conducted at only one site, in one year, not within the native range of *A. thaliana*, and at a latitude south those from which the accessions in the association analyses were collected, the mechanisms of selection affecting *FRI* that we observed may not be the same as those that generated patterns of *FRI* diversity across Europe. Nevertheless, our results suggest possible explanations for some patterns of *FRI FLC* genotype diversity and lead to testable predictions. Functional *FRI FLC*^{A} genotypes were favored in the fall cohort, and functional *FRI FLC*^{B} genotypes were selected against in the spring cohort, providing a potential selective mechanism for the over- and underrepresentation of these allele combinations in natural European populations (Caicedo et al. 2004). In addition, because null *FRI* alleles were favored only in the spring cohort, our results suggest that a climate permitting a successful spring cohort is necessary for the success of null alleles. This may explain why *FRI* null alleles are common in the mild oceanic climate of northwestern Europe (Hagenblad et al. 2004; see fig. A1). Although the geographical distribution of a spring- or summer-germinating generation has not been well surveyed, spring and summer germinants have been observed in England and western continental Europe but not in Mediterranean regions (Thompson 1994; C. Alonso-Blanco and M. Koornneef, personal communications).

Our results do not suggest an explanation for why there is a cline in *FLC* variation and why *FLC*^{B} is common in northern Europe (Caicedo et al. 2004), because we found no selective advantage for the *FLC*^{B} allele under our experimental conditions. The *FLC*^{B} allele may confer an advantage in an environment we did not test, such as at high latitudes, in summer-germinating generations, or in genetic backgrounds found in northern Europe that were not well represented in our sample. Alternatively, this cline may be due to historical population structure or linkage to another gene affected by clinal selection.

While our results show that there were fitness associations with *FRI* and *FLC* in our experiment that would produce large changes in allele frequencies within a generation, we cannot rule out the possibility that this selection was indirect and due to the phenotypic effects of other genes. As in all association studies, care must be taken in attributing the cause of an association to a candidate gene (Page et al. 2003). It is possible that loci closely physically linked to *FRI* are responsible for the associations, and the high linkage disequilibrium surrounding the *FRI* locus makes it unlikely that we could narrow down the region associated with the fitness effects to less than 30 kb with flanking markers (Hagenblad et al. 2004). In addition, we cannot rule out that cryptic population structure unaccounted for by our data caused the associations (e.g., Campbell et al. 2005). In particular, some effects of *FRI* may not be dependent on the *FLC* allele present but may instead depend on alleles present at loci in linkage disequilibrium with *FLC*. Nevertheless, several lines of evidence suggest that *FRI* is the cause of the fitness associations. First, in each season, two evolutionarily independent *FRI* null alleles had similar fitness associations. Neither of these *FRI* null alleles are strongly associated with haplotypes across the genome (Aranzana et al. 2006), making it even less likely that their associations with fitness are due to cryptic population structure.

Another reason to believe that *FRI* is the cause of the fitness association in the fall cohort is that we identified a mechanism for the effect of *FRI* on winter survival. *FRI* genotype and winter survival were both associated with time to bolting, a trait that *FRI* is well known to affect. However, *FRI* delayed bolting in both *FLC* backgrounds, suggesting that a survival effect of *FRI* should not be dependent on *FLC* genotype. This suggests that there may be other loci in the FLC^B background that are responsible for the low winter survivorship of accessions with *FRI* alleles. Alternatively, in accessions with functional *FRI* alleles, FLC^B alleles might reduce winter survival via effects on traits we did not measure.

In the spring cohort, the *FRI* × *FLC* epistatic association with total seed mass could not be explained by selection on bolting time. Instead, *FRI FLC* genotype and seed production were both associated with rosette growth. While rosette growth has not previously been connected with *FRI*, *FLC* integrates signals from the autonomous pathway, which triggers flowering based, in part, on plant size (Boss et al. 2004). This suggests that *FRI FLC* genotypes may differ in interactions with the autonomous pathway and that effects of *FLC* variation on rosette growth might be worth investigating. Functional *FRI* and *FLC* alleles are also associated with greater water use efficiency in near-isogenic lines (McKay et al. 2003). It is possible that the *FRI* FLC^B genotype was associated with especially high water use efficiency, which may have conferred slower growth and thus proved maladaptive in the wet spring environment.

Our results for time to bolting in the spring, in which functional *FRI* alleles did not delay time to bolting in the FLC^B background, as they did in the FLC^A background, suggest that at least a significant subset of the FLC^B alleles may be weak. By weak, we mean that an allele does not delay bolting time, when combined with a functional *FRI* allele under nonvernalized conditions, to as great an extent as other, strong alleles. Other studies have shown and/or suggest that the FLC^B alleles in a number of accessions are weak (Bd-0 [Lempe et al. 2005], Per-1 [Shindo et al. 2005], Kas-1 [El-Lithy et al. 2006], Kondara [Michaels et al. 2003; El-Lithy et al. 2006], Shahdara [Gazzani et al. 2003; Shindo et al. 2005], and Wa-1 [Shindo et al. 2005; Werner et al. 2005*a*]). It is possible that the sequence difference that distinguishes FLC^B from FLC^A does not cause a weak allele but instead that there are one or more common variants within the FLC^B haplogroup that result in weak alleles.

Even though the *FRI* locus can explain up to 70% of variation in flowering time under certain conditions (Shindo et al. 2005), it is not surprising that *FRI* and *FLC* explained only 5.4% of winter survival variation and 8% of variation in seed production in the spring. *FRI* functional and *FRI* null lines differ most strongly in flowering time in nonvernalized conditions and much less so in vernalized conditions (Lempe et al. 2005; Shindo et al. 2005). In the fall cohort, plants received vernalization via the winter, and in the spring cohort, plants probably received some vernalization due to a 2-week cold treatment as seeds and during chilly early-spring nights. The fact that a small percentage of markers showed more significant associations than *FRI* within *FLC* backgrounds is also not surprising, for two reasons. First, we expect there to be some other loci with strong effects on fitness traits under field conditions. Second, some markers may have strong associations due to population structure not accounted for by the controls for population structure (Aranzana et al. 2006).

Spurious candidate gene associations can be caused by cryptic population structure (Knowler et al. 1988; Hoggart et al. 2003). Recent studies suggest that *A. thaliana* has substantial population structure (Nordborg et al. 2005; Schmid et al. 2006) and that this structure can affect associations (Aranzana et al. 2006). Our study employed two methods to mitigate this problem: using inferred-ancestry estimates and restricting the geographical origin of our samples. Estimates of inferred ancestry have been used in human and maize associations (Thornsberry et al. 2001; Hoggart et al. 2003; Wilson et al. 2004) and a recent *A. thaliana* study (Aranzana et al. 2006), but this is the first study employing this approach for fitness traits under field conditions and the first using it in concert with a geographical restriction. Others have suggested that each of these methods could be important for reducing the rate of false positives (Aranzana et al. 2006). Our study suggests that in addition, these methods may increase power and enable the detection of associations that would otherwise be missed. The geographical restriction had a large effect because it removed samples that obscured associations, possibly because of the confounding of geographic origin, genotype, and trait values. The inferred-ancestry estimates had an effect because they removed a small amount of variation between different genetic backgrounds that was not attributable to the candidate genes. These inferred-ancestry estimates are not intended to accurately represent the structure of the genetic diversity of *A. thaliana*, because *A. thaliana* exhibits genetic isolation by distance rather than discrete subpopulations (Sharbel et al. 2000; Nordborg et al. 2005; Schmid et al. 2006), but nevertheless may be a useful construct for excluding some of the effects of genetic background.

Our results suggest that seasonally variable selection and epistasis were critical to the evolution of *FRI* null alleles. Furthermore, our results suggest that the combined effects of environmental heterogeneity and epistasis may be important for maintaining genetic variation underlying flowering time. Unlike in some species, as-

sociations in *A. thaliana* can be tested experimentally because of the genetic tools available and because *A. thaliana* can be grown in experiments in ecologically realistic conditions. Further experiments are underway to test whether *FRI* and *FLC* are the causative agents of the fitness effects we observed. However, taken together with several other recent studies of fitness in recombinant inbred lines (Weinig et al. 2003; Kroymann and Mitchell-Olds 2005; Malmberg et al. 2005), our results suggest that epistasis and selection in heterogeneous environments may be crucial for the maintenance of genetic polymorphism in *A. thaliana* and other species.

Acknowledgments

We thank E. de Moor, J. Plaut, N. Reese, and B. Singh for their assistance in the field, E. von Wettberg for making the map, and T. Altmann and O. Törjék for the SNP genotyping. The work was supported by National Science Foundation grants DEB-9976997 and EF-0425759.

APPENDIX A

Methods

Methods for Determining Total Seed Mass

On senescence, plants were harvested and the fruits were counted. To estimate seed mass per fruit, as the plants senesced, five fruits were collected from each plant from representative positions based on the distribution of fruits on the branches. Seed from the collected fruits was weighed, and total seed mass per plant was calculated by multiplying seed mass per fruit by number of fruits. Nondehisced fruits were not collected in time to measure seed mass for 5.5% of fruit-producing plants in the fall and 5.6% in the spring. For these plants, seed mass was approximated on the basis of the relationships of block and log (total fruit number) with seed mass per fruit. We calculated the parameters for an ANCOVA model, seed mass per fruit $= a + b_i \times \text{block}_i + c_i \times \log(\text{number of fruits})$, and used these parameters to estimate seed mass per fruit for those with missing values.

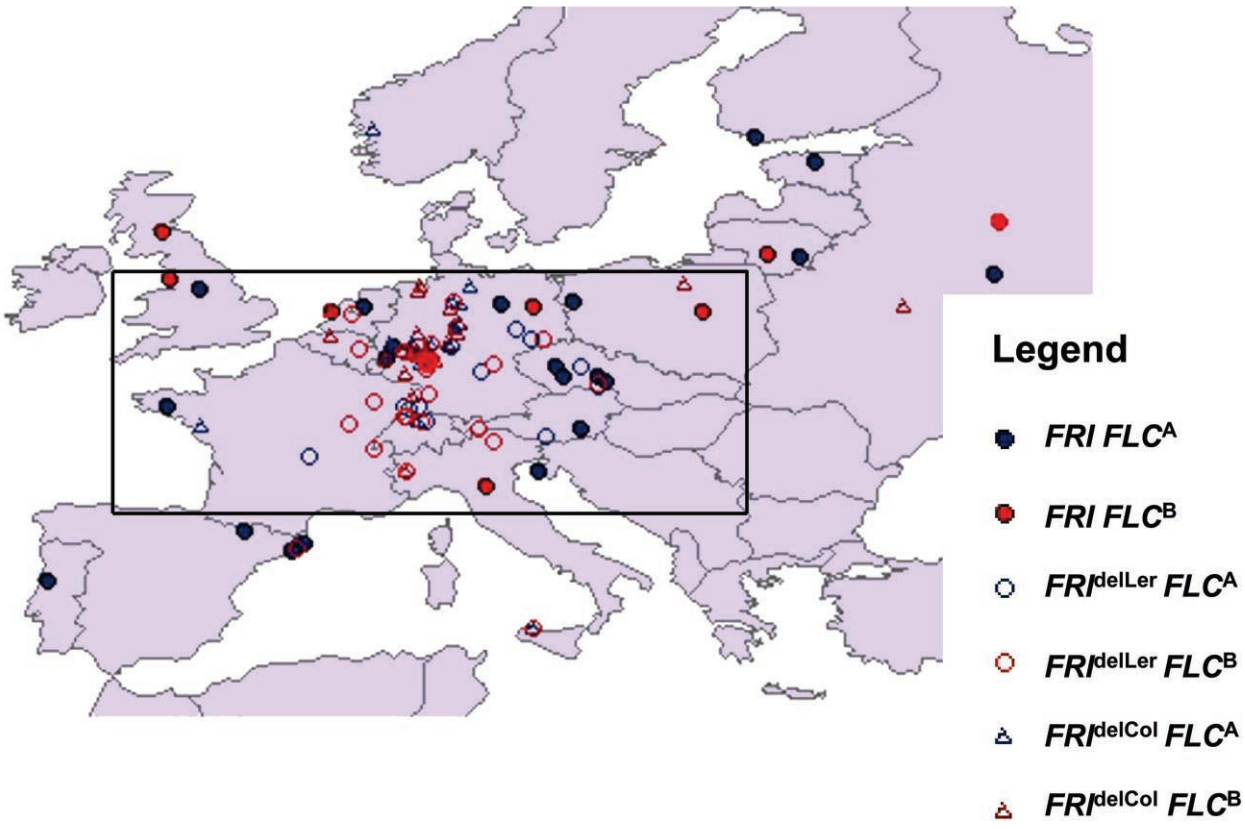

Figure A1: Map of accessions' origins with their *FRI FLC* genotypes. The box encloses the geographically restricted region used in association analyses.

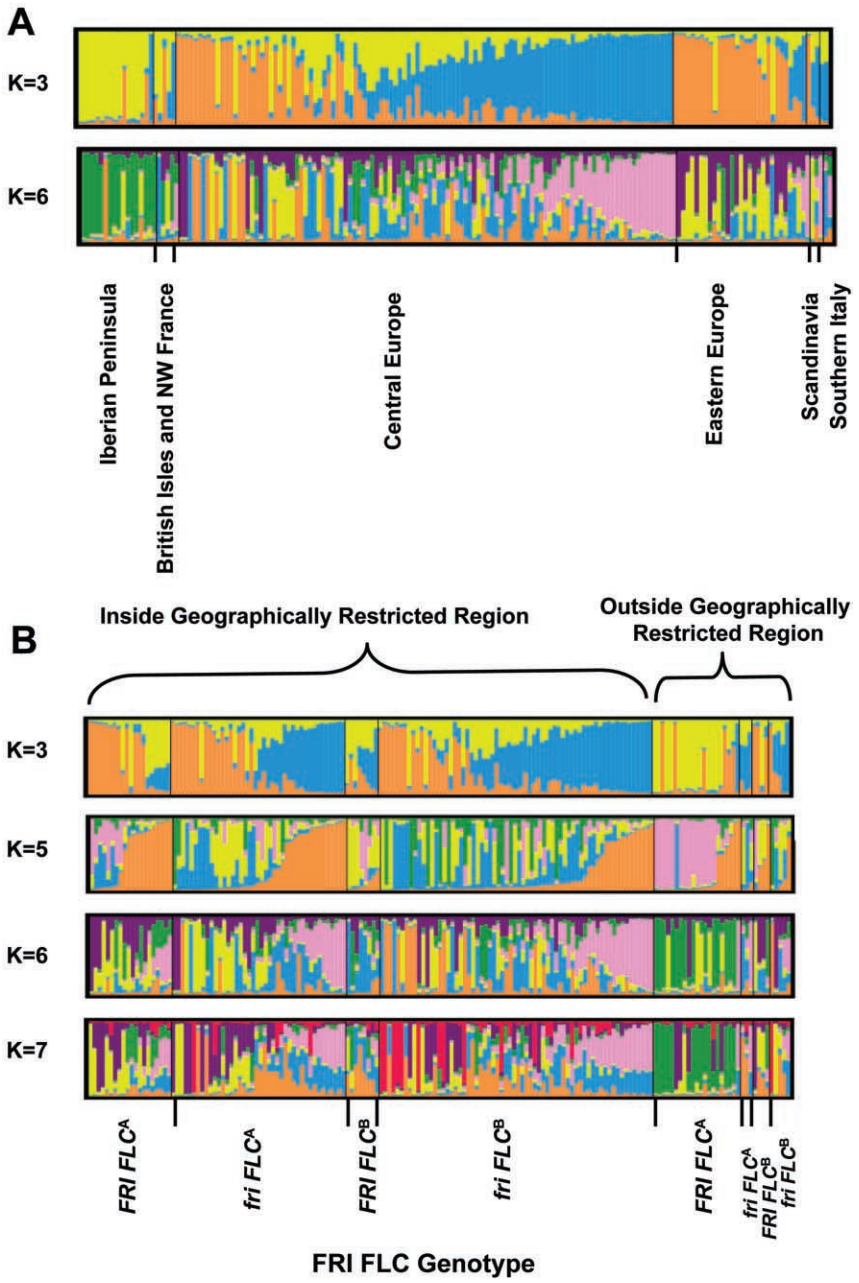

Figure A2: *Distruct* images for inferred ancestry of the accessions. Each color represents an ancestral population, and the relative heights of the colors indicate the proportions of ancestry for each accession. *A*, Accessions grouped by geographical origin, for number of populations $K = 3$ and 6. *B*, Accessions grouped by *FRI FLC* genotype and whether the accession was included in our geographically restricted set used for associations, for $K = 3$, 5, 6, and 7. For $K = 3$ and $K = 6$, the colors correspond to the same ancestral populations in *A* and *B*. The colors in the $K = 6$ plots correspond to the following ancestries in "Results" and the zip archive: orange = population 1; blue = population 2; yellow = population 3; pink = population 4; green = population 5; purple = population 6. In both plots, accessions are sorted by ancestry within each category, and the order of the accessions differs for different *K*. See Schmid et al. (2006) for additional results about population structure and geographical origin using a larger, overlapping SNP data set.

Table A1: Estimated log likelihoods of the data, given the number of populations (K) and probabilities for the number of populations

K	log $P(D\|K)$[a]	$P(K\|D)$[b]
1	−6,399.1	3.20×10^{-179}
2	−6,257.0	1.65×10^{-117}
3	−6,131.5	5.27×10^{-63}
4	−6,085.4	5.54×10^{-43}
5	−6,000.4	4.55×10^{-6}
6	−5,988.1	1.00
7	−6,063.7	1.47×10^{-33}

[a] Log-likelihood values are the median values observed in five runs for each K.

[b] Calculated as in Pritchard et al. (2000*b*); assumes a uniform prior for K.

Methods for Estimating Population Ancestry

Because *Arabidopsis thaliana* is highly selfing and because no options for partial selfing are currently available in *structure*, we entered the data as for a haploid organism (J. Pritchard, personal communication), with missing data for rare heterozygous loci. The program was run with admixture and correlated gene frequencies between populations using a burn-in period of 30,000 runs and 40,000 repetitions for parameter estimation. Results presented by Schmid et al. (2006) suggest that ascertainment bias does not affect the estimation of the number of populations with this SNP data set.

These inferred-ancestry results are similar to those of Schmid et al. (2006), which used the same SNP data set with a larger set of accessions, and appear to be consistent with population structure results reported by Nordborg et al. (2005). There are some deviations between our population structure results and those of Schmid et al. (2006) because of that study's inclusion of Asian accessions. Asian samples were not included in population analyses here because Asian population ancestry is very low in the set of European accessions used in our association analyses. Nordborg et al. (2005) include accessions that span a larger geographic range than those in our study and consequently include some population ancestries not represented in our sample. Our study also assigns a greater number of ancestral populations to central Europe.

APPENDIX B

Additional Results about the Effects of Inferred Ancestry and Geographic Restrictions on Associations

Additional Ancestry Restriction

The ancestry composition of the *fri* FLC^B accessions (deletion null *FRI* accessions with FLC^B) is not completely represented in the small set of *FRI* FLC^B accessions. This can be seen in figure A2*B*, in the $K = 6$ image sorted by genotype. In particular, there are three ancestries that are present in a number of the *fri* FLC^B accessions at high proportions that are not present in any of the *FRI* FLC^B accessions in high proportions.

To investigate whether these differences in ancestries may have had an effect on the association we observed in FLC^B for seed production, we performed an analysis on a data set in which we eliminated all accessions with greater than 50% ancestry from the three populations not well represented among the *FRI* FLC^B accessions. This reduced the number of accessions from 136 to 81.

Restricting the data set in this way did not affect the associations we observed in the FLC^B background. In the spring cohort, *FRI* FLC^B accessions still had lower seed production than *fri* FLC^B accessions (in FLC^B, means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 4.76$, df $= 1, 70$, $P = .033$; 2.7% of SNPs [2 of 75] had a higher F statistic for a means contrast in the FLC^B background). For winter survival in the fall cohort and for time to bolting in the spring cohort, there were still no differences (for winter survival in FLC^B, means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 1.02$, df $= 1, 70$, $P = .31$; for time to bolting in spring in FLC^B, means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 0.062$, df $= 1, 70$, $P = .80$).

Additional Geographical Restriction

We reduced the set of accessions for association analyses to those from central Europe, eliminating five accessions west of 2°E longitude and two accessions east of 17°E longitude; these cutoffs were chosen based on geographical gaps in our sample (see fig. A1). This reduced the data set from 136 to 129 accessions and the number of *FRI* FLC^B accessions from eight to six.

With this new data set, the directions of associations were the same, and means contrasts within the *FLC* backgrounds were still statistically significant or nearly so. *FRI* functionality × *FLC* epistasis was no longer detectable for winter survival (*FRI* functionality × *FLC*: $F = 1.35$, df $= 1, 120$, $P = .25$) or spring seed production (*FRI* functionality × *FLC*: $F = 2.42$, df $= 1, 120$, $P = .12$), probably because of the loss of power from the low number of *FRI* FLC^B accessions.

For winter survival, in the FLC^A background, we observed a significant effect of *FRI* functionality, as in the larger data set (means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 5.34$, df $= 1, 118$, $P = .023$; 6.4% of SNPs [5 of 78] had a higher F statistic for associations with winter survival). In the FLC^B background, we observed no effect, as in the larger data set (means contrast

Table B1: Effect of geographical restrictions and single nucleotide polymorphism ancestry covariates on the association of *FRI* functionality with time to bolting in Stinchcombe et al. (2004)

Line	Geographical restriction	Ancestry covariates	N	F[a]	df	P	Days to bolting (Julian date)[b]		
							FRI	FRI^{delLer}	FRI^{delCol}
1	44°–54°N latitude, west of 22°E longitude	Yes	38	4.43	1, 27	.044	167.6 ± 2.8	165.1 ± 2.8	155.5 ± 3.4
2	44°–54°N latitude, west of 22°E longitude	No	38	.56	1, 32	.46	164.1 ± 2.7	159.7 ± 2.3	163.7 ± 2.5
3	No Spanish accessions	Yes	47	2.7	1, 36	.11	166.4 ± 2.3	162.5 ± 2.5	160.8 ± 2.7
4	No Spanish accessions	No	47	.36	1, 41	.55	164.0 ± 2.1	159.7 ± 2.2	165.3 ± 2.1
5	None	Yes	52	.26	1, 41	.62	161.6 ± 2.8	162.0 ± 3.4	157.4 ± 3.8
6	None	No	52	4.16	1, 46	.047	155.8 ± 2.4	159.7 ± 3.2	165.3 ± 3.0

[a] F statistics and P values are from means contrasts of *FRI* versus FRI^{delLer} and FRI^{delCol} in FLC^{A}. Statistics for FLC^{B} genotypes are not shown because of very low sample sizes.

[b] Values are least squares means ± SE.

of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 0.10$, df = 1, 118, $P = .75$).

For total seed mass per plant in the spring cohort, in the FLC^{A} background, there was no effect of *FRI* functionality in the reduced data set, as in the larger data set (means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 0.008$, df = 1, 118, $P = .93$). In the FLC^{B} background, where a significant difference was detected in the larger data set, the effect was marginally significant in the reduced data set (means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 3.77$, df = 1, 118, $P = .055$; 6.4% of SNPs [5 of 78] had a higher F statistic for associations with seed production).

For time to bolting in the spring cohort, in the FLC^{A} background, there was a significant effect of *FRI* functionality, as in the larger data set (means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 13.4$, df = 1, 118, $P = .0004$). In the FLC^{B} background, no difference was detected, as in the larger data set (means contrast of *FRI* vs. FRI^{delLer} and FRI^{delCol} alleles: $F = 0.074$, df = 1, 118, $P = .79$). No SNPs (0 of 78) had a higher F statistic for associations in either *FLC* background. There was still a significant *FRI* functionality × *FLC* interaction ($F = 4.35$, df = 1, 120, $P = .039$).

Effects of Inferred Ancestry Covariates and Geographical Restrictions on Associations with Time-to-Bolting Data from Stinchcombe et al. (2004)

We investigated the effects of different treatments of population structure on associations with time-to-bolting data from Stinchcombe et al. (2004) and Caicedo et al. (2004; table B1). Because we had only inferred-ancestry estimates for accessions in our experiment, we used a set of accessions that included only those that overlapped between the experiments. Without our controls for population structure, we obtained a result similar to that reported in Caicedo et al. (table B1, line 6): accessions with *FRI* bolted more quickly than accessions with *fri* in the FLC^{A} background. With both our geographical restriction and inferred-ancestry covariates in the model, we found that accessions with *FRI* bolted later than accessions with *fri* in the FLC^{A} background (table B1, line 1). Both our geographical restriction and the inferred-ancestry covariates were necessary for this association (table B1, lines 1, 2, and 5).

Additional analyses indicate that a major factor contributing to the different results with and without our controls for population structure is the inclusion of accessions from Spain (table B1, lines 1, 3, and 5). Spanish accessions tended to bolt early (on average 24 days earlier than accessions from other regions; one-way ANOVA: $F = 51.8$, df = 1, 50, $P < .0001$), had little diversity in *FRI FLC* genotype (15 of the 16 Spanish accessions had the *FRI* FLC^{A} genotype), and were distinct in their population ancestry, as determined by the SNP markers (13 of the 16 Spanish accessions had greater than 75% ancestry from one estimated population, whereas no accessions from other locations have such high ancestry associated with this population; see fig. A2*A*). Consequently, we cannot determine whether early bolting in the Spanish accessions is caused by the *FRI* FLC^{A} genotype or by the genotype at other background loci (see Hagenblad et al. 2004 for a similar example).

Caicedo et al. (2004) attempted to account for population structure using an AFLP data set (Sharbel et al. 2000; Olsen et al. 2004) and identified an unstratified population sample that included the Spanish accessions, a conclusion supported by a neighbor-joining analysis of the data. The population structure using SNPs in this study may be more accurate because the results concur with those of another recent study also based on SNPs (Nordborg et al. 2005). It is unclear why the AFLP and SNP marker data sets provide contrasting results about population structure.

This may arise from differences in the number of markers (79 AFLPs vs. 115 SNPs), the number of accessions in the data sets (104 for the analysis with AFLPs vs. 169 in this study), the mutational dynamics of these marker types, the degree of ascertainment bias in the case of the SNP data set used in our study, and/or the informativeness of the markers. In addition, we cannot rule out the possibility that *FRI FLC* genotypes have different effects in genetic backgrounds prevalent in other regions, and our conclusions with the controls for population structure apply only to the restricted geographical range of our study.

No Effect of Inferred-Ancestry Covariates on the Latitudinal Cline in Time to Bolting Reported by Stinchcombe et al. (2004)

We examined whether population structure, as determined from the SNP data, could explain the cline in time to bolting in Stinchcombe et al. (2004). Caicedo et al. (2004) used the same time-to-bolting data as Stinchcombe et al. When the inferred-ancestry estimates were included in the analysis, the latitudinal cline and *FRI* × latitude interaction for bolting time remained ($N = 52$ accessions; latitude main effect: $\mathrm{df} = 1, 43$, $P = .006$; *FRI* functionality × latitude: $\mathrm{df} = 1, 43$, $P = .005$), suggesting that the cline in accessions with *FRI* functional alleles is not due to population structure.

Literature Cited

Aranzana, M. J., S. Kim, K. Zhao, E. Bakker, M. Horton, K. Jakob, C. Lister, et al. 2006. Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. Public Library of Science Genetics 1:e60, doi: 10.1371/journal.pgen.0010060.

Bakker, E. G., E. A. Stahl, C. Toomajian, M. Nordborg, M. Kreitman, and J. Bergelson. 2006. Distribution of genetic variation within and among local populations of *Arabidopsis thaliana* over its species range. Molecular Ecology 15:1405–1418.

Boss, P. K., R. M. Bastow, J. S. Mylne, and C. Dean. 2004. Multiple pathways in the decision to flower: enabling, promoting, and resetting. Plant Cell 16:S18–S31.

Caicedo, A. L., J. R. Stinchcombe, K. M. Olsen, J. Schmitt, and M. D. Purugganan. 2004. Epistatic interaction between *Arabidopsis FRI* and *FLC* flowering time genes generates a latitudinal cline in a life history trait. Proceedings of the National Academy of Sciences of the USA 101:15670–15675.

Campbell, C. D., E. L. Ogburn, K. L. Lunetta, H. N. Lyon, M. L. Freedman, L. C. Groop, D. Altshuler, K. J. Ardlie, and J. N. Hirschhorn. 2005. Demonstrating stratification in a European American population. Nature Genetics 37:868–872.

Cardon, L. R., and L. J. Palmer. 2003. Population stratification and spurious allelic association. Lancet 361:598–604.

Donohue, K. 2002. Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. Ecology 83:1006–1016.

Donohue, K., L. Dorn, C. Griffith, E. Kim, A. Aguilera, C. R. Polisetty, and J. Schmitt. 2005. The evolutionary ecology of seed germination of *Arabidopsis thaliana*: variable natural selection on germination timing. Evolution 59:758–770.

Eanes, W. F. 1999. Analysis of selection on enzyme polymorphisms. Annual Review of Ecology and Systematics 30:301–326.

El-Assal, S. E.-D., C. Alonso-Blanco, A. J. M. Peeters, V. Raz, and M. Koornneef. 2001. A QTL for flowering time in *Arabidopsis* reveals a novel allele of *CRY2*. Nature Genetics 29:435–440.

El-Lithy, M. E., L. Bentsink, C. J. Hanhart, G. J. Ruys, D. Rovito, J. L. M. Broekhof, H. J. A. van der Poel, M. J. T. van Eijk, D. Vreugdenhil, and M. Koornneef. 2006. New *Arabidopsis* recombinant inbred line populations genotyped using SNPWave and their use for mapping flowering-time QTLs. Genetics 172:1867–1876.

Ellner, S., and N. G. Hairston. 1994. Role of overlapping generations in maintaining genetic variation in a fluctuating environment. American Naturalist 143:403–417.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon, Oxford.

Gazzani, S., A. R. Gendall, C. Lister, and C. Dean. 2003. Analysis of the molecular basis of flowering time variation in *Arabidopsis* accessions. Plant Physiology 132:1107–1114.

Gillespie, J. H., and M. Turelli. 1989. Genotype-environment interactions and the maintenance of polygenic variation. Genetics 121: 129–138.

Gimelfarb, A. 1989. Genotypic variation for a quantitative character maintained under stabilizing selection without mutations: epistasis. Genetics 123:217–227.

Glazier, A. M., J. H. Nadeau, and T. J. Aitman. 2002. Finding genes that underlie complex traits. Science 298:2345–2349.

Griffith, C., E. Kim, and K. Donohue. 2004. Life-history variation and adaptation in the historically mobile plant *Arabidopsis thaliana* (Brassicaceae) in North America. American Journal of Botany 91: 837–849.

Hagenblad, J., C. Tang, J. Molitor, J. Werner, K. Zhao, H. Zheng, P. Marjoram, D. Weigel, and M. Nordborg. 2004. Haplotype structure and phenotypic associations in the chromosomal regions surrounding two *Arabidopsis thaliana* flowering time loci. Genetics 168:1627–1638.

Hanson, M. A., B. S. Gaut, A. O. Stec, S. I. Fuerstenberg, M. M. Goodman, E. H. Coe, and J. F. Doebley. 1996. Evolution of anthocyanin biosynthesis in maize kernels: the role of regulatory and enzymatic loci. Genetics 143:1395–1407.

Hoekstra, H. E., K. E. Drumm, and M. W. Nachman. 2004. Ecological genetics of adaptive color polymorphism in pocket mice: geographic variation in selected and neutral genes. Evolution 58:1329–1341.

Hoggart, C. J., E. J. Parra, M. D. Shriver, C. Bonilla, R. A. Kittles, D. G. Clayton, and P. M. McKeigue. 2003. Control of confounding of genetic associations in stratified populations. American Journal of Human Genetics 72:1492–1504.

Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino, and C. Dean. 2000. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290: 344–347.

Knowler, W. C., R. C. Williams, D. J. Pettitt, and A. G. Steinberg. 1988. Gm3-5,13,14 and type-2 diabetes-mellitus: an association in American Indians with genetic admixture. American Journal of Human Genetics 43:520–526.

Kroymann, J., and T. Mitchell-Olds. 2005. Epistasis and balanced polymorphism influencing complex trait variation. Nature 435: 95–98.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated traits. Evolution 37:1210–1226.

Le Corre, V. 2005. Variation at two flowering time genes within and among populations of *Arabidopsis thaliana*: comparison with markers and traits. Molecular Ecology 14:4181–4192.

Le Corre, V., F. Roux, and X. Reboud. 2002. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time. Molecular Biology and Evolution 19:1261–1271.

Leips, J., and T. F. C. Mackay. 2000. Quantitative trait loci for life span in *Drosophila melanogaster*: interactions with genetic background and larval density. Genetics 155:1773–1788.

Lempe, J., S. Balasubramanian, S. Sureshkumar, A. Singh, M. Schmid, and D. Weigel. 2005. Diversity of flowering responses in wild *Arabidopsis thaliana* strains. Public Library of Science Genetics 1:e6, doi:10.1371/journal.pgen.0010006.

Levene, H. 1953. Genetic equilibrium when more than one ecological niche is available. American Naturalist 87:331–333.

Long, A. D., R. F. Lyman, A. H. Morgan, C. H. Langley, and T. F. C. Mackay. 2000. Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the achaete-scute complex are associated with variation in bristle number in *Drosophila melanogaster*. Genetics 154:1255–1269.

Malmberg, R. L., S. Held, A. Waits, and R. Mauricio. 2005. Epistasis for fitness-related quantitative triats in *Arabidopsis thaliana* grown in the field and in the greenhouse. Genetics 171:2013–2027.

Mauricio, R., E. A. Stahl, T. Korves, D. C. Tian, M. Kreitman, and J. Bergelson. 2003. Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana*. Genetics 163: 735–746.

McKay, J. K., J. H. Richards, and T. Mitchell-Olds. 2003. Genetics of drought adaptation in *Arabidopsis thaliana*. I. Pleiotropy contributes to genetic correlations among ecological traits. Molecular Ecology 12:1137–1151.

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

Michaels, S. D., Y. He, K. C. Scortecci, and R. M. Amasino. 2003. Attenuation of *FLOWERING LOCUS C* activity as a mechanism for the evolution of summer-annual flowering behavior in *Arabidopsis*. Proceedings of the National Academy of Sciences of the USA 100:10102–10107.

Napp-Zinn, K. 1987. Vernalization: environmental and genetic regulation. Pages 123–132 *in* J. G. Atherton, ed. Manipulation of flowering. Butterworths, London.

Nordborg, M., T. T. Hu, Y. Ishino, J. Jharveri, C. Toomajian, H. Zheng, E. Bakker, et al. 2005. The pattern of polymorphism in *Arabidopsis thaliana*. Public Library of Science Biology 3:e196, doi: 10.1371/journal.pbio.0030196.

Olsen, K. M., A. Womack, A. R. Garrett, J. I. Suddith, and M. D. Purugganan. 2002. Contrasting evolutionary forces in the *Arabidopsis thaliana* floral developmental pathway. Genetics 160:1641–1650.

Olsen, K. M., S. S. Halldorsdottir, J. R. Stinchcombe, C. Weinig, J. Schmitt, and M. D. Purugganan. 2004. Linkage disequilibrium mapping of *Arabidopsis CRY2* flowering time alleles. Genetics 167: 1361–1369.

OMAF. 2002. Spring and winter canola *in* T. Baute, ed. Agronomy guide for field crops. Ontario Ministry of Food and Agriculture, Toronto.

Page, G. P., V. George, R. C. Go, P. Z. Page, and D. B. Allison. 2003. "Are we there yet?": deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. American Journal of Human Genetics 73:711–719.

Peripato, A. C., R. A. De Brito, S. R. Matioli, L. S. Pletscher, T. Vaughn, and J. M. Cheverud. 2004. Epistasis affecting litter size in mice. Journal of Evolutionary Biology 17:593–602.

Pigliucci, M., and E. T. Marlow. 2001. Differentiation for flowering time and phenotypic integration in *Arabidopsis thaliana* in response to season length and vernalization. Oecologia (Berlin) 127: 501–508.

Pritchard, J. K., M. Stephens, N. A. Rosenberg, and P. Donnelly. 2000*a*. Association mapping in structured populations. American Journal of Human Genetics 67:170–181.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000*b*. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Rausher, M. D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. Evolution 46:616–626.

Routman, E. J., and J. M. Cheverud. 1997. Gene effects on a quantitative trait: two-locus epistatic effects measured at microsatellite markers and at estimated QTL. Evolution 51:1654–1662.

Schmid, K. J., S. Ramos-Onsins, H. Ringys-Beckstein, B. Weisshaar, and T. Mitchell-Olds. 2005. A multilocus sequence survey in *Arabidopsis thaliana* reveals a genome-wide departure from a neutral model of DNA sequence polymorphism. Genetics 169:1601–1615.

Schmid, K. J., O. Törjék, R. Meyer, H. Schmuths, M. Hoffmann, and T. Altmann. 2006. Evidence for a large-scale population structure of *Arabidopsis thaliana* from genome-wide SNP markers. Theoretical and Applied Genetics 112:1104–1114.

Sharbel, T. F., B. Haubold, and T. Mitchell-Olds. 2000. Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. Molecular Ecology 9:2109–2118.

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

Shook, D. R., and T. E. Johnson. 1999. Quantitative trait loci affecting survival and fertility-related traits in *Caenorhabditis elegans* show genotype-environment interactions, pleiotropy and epistasis. Genetics 153:1233–1243.

Simpson, G. G., and C. Dean. 2002. Flowering: *Arabidopsis*, the rosetta stone of flowering time? Science 296:285–289.

Stahl, E. A., G. Dwyer, R. Mauricio, M. Kreitman, and J. Bergelson. 1999. Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. Nature 400:667–671.

Stenoien, H. K., C. B. Fenster, A. Tonteri, and O. Savolainen. 2005. Genetic variability in natural populations of *Arabidopsis thaliana* in northern Europe. Molecular Ecology 14:137–148.

Stinchcombe, J. R., M. T. Rutter, D. S. Burdick, P. Tiffin, M. D. Rausher, and R. Mauricio. 2002. Testing for environmentally induced bias in phenotypic estimates of natural selection: theory and practice. American Naturalist 160:511–523.

Stinchcombe, J. R., C. Weinig, M. Ungerer, K. M. Olsen, C. Mays, S. S. Halldorsdottir, M. D. Purugganan, et al. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the

flowering time gene *FRIGIDA*. Proceedings of the National Academy of Sciences of the USA 101:4712–4717.

Thompson, L. 1994. The spatiotemporal effects of nitrogen and litter on the population dynamics of *Arabidopsis thaliana*. Journal of Ecology 82:63–68.

Thornsberry, J. M., M. M. Goodman, J. Doebley, S. Kresovich, D. Nielsen, and E. S. Buckler. 2001. *Dwarf8* polymorphisms associate with variation in flowering time. Nature Genetics 28:286–289.

Tian, D., H. Araki, E. Stahl, J. Bergelson, and M. Kreitman. 2002. Signature of balancing selection in *Arabidopsis*. Proceedings of the National Academy of Sciences of the USA 99:11525–11530.

Tian, D., M. B. Traw, J. Q. Chen, M. Kreitman, and J. Bergelson. 2003. Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. Nature 423:74–77.

Toomajian, C., T. T. Hu, M. J. Aranzana, C. Lister, C. Tang, H. Zheng, K. Zhao, P. Calabrese, C. Dean, and M. Nordborg. 2006. A nonparametric test reveals selection for rapid flowering in the *Arabidopsis* genome. Public Library of Science Biology 4:e137, doi: 10.1371/journal.pbio.0040137.

Voight, B. F., S. Kudaravalli, X. Q. Wen, and J. K. Pritchard. 2006. A map of recent positive selection in the human genome. Public Library of Science Biology 4:e72, doi:10.1371/journal.pbio.0040072.

Wade, M. J., and C. J. Goodnight. 1998. The theories of Fisher and Wright in the context of metapopulations: when nature does many small experiments. Evolution 52:1537–1553.

Watt, W. B. 1977. Adaptation at specific loci. I. Natural selection on phosphoglucose isomerase of colias butterflies: biochemical and population aspects. Genetics 87:177–194.

Weinig, C., M. C. Ungerer, L. A. Dorn, N. C. Kane, Y. Toyonaga, S. S. Halldorsdottir, T. F. C. Mackay, M. D. Purugganan, and J. Schmitt. 2002. Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. Genetics 162: 1875–1884.

Weinig, C., L. A. Dorn, N. C. Kane, Z. M. German, S. S. Halldorsdottir, M. C. Ungerer, Y. Toyonaga, T. F. C. Mackay, M. D. Purugganan, and J. Schmitt. 2003. Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*. Genetics 165: 321–329.

Weinreich, D. M., R. A. Watson, and L. Chao. 2005. Sign epistasis and genetic constraint on evolutionary trajectories. Evolution 59: 1165–1174.

Werner, J. D., J. O. Borevitz, N. H. Uhlenhaut, J. R. Ecker, J. Chory, and D. Weigel. 2005*a*. *FRIGIDA*-independent variation in flowering time of natural *Arabidopsis thaliana* accessions. Genetics 170: 1197–1207.

Werner, J. D., J. O. Borevitz, N. Warthmann, G. T. Trainer, J. R. Ecker, J. Chory, and D. Weigel. 2005*b*. Quantitative trait locus mapping and DNA array hybridization identify an *FLM* deletion as a cause for natural flowering-time variation. Proceedings of the National Academy of Sciences of the USA 102:2460–2465.

Whitlock, M. C., P. C. Phillips, F. B. G. Moore, and S. J. Tonsor. 1995. Multiple fitness peaks and epistasis. Annual Review of Ecology and Systematics 26:601–629.

Wilson, L. M., S. R. Whitt, A. M. Ibanez, T. R. Rocheford, M. M. Goodman, and E. S. Buckler. 2004. Dissection of maize kernel composition and starch production by candidate gene association. Plant Cell 16:2719–2733.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16: 97–159.

Wright, S. I., and B. S. Gaut. 2005. Molecular population genetics and the search for adaptive evolution in plants. Molecular Biology and Evolution 22:506–519.

Associate Editor: George W. Gilchrist
Editor: Michael C. Whitlock