# Independent *FLC* Mutations as Causes of Flowering-Time Variation in *Arabidopsis thaliana* and *Capsella rubella*

**Ya-Long Guo,*,† Marco Todesco,* Jörg Hagmann,* Sandip Das,*,[1] and Detlef Weigel*,[2]**
*Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany, and †State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

**ABSTRACT** *Capsella rubella* is an inbreeding annual forb closely related to *Arabidopsis thaliana*, a model species widely used for studying natural variation in adaptive traits such as flowering time. Although mutations in dozens of genes can affect flowering of *A. thaliana* in the laboratory, only a handful of such genes vary in natural populations. Chief among these are *FRIGIDA* (*FRI*) and *FLOWERING LOCUS C* (*FLC*). Common and rare *FRI* mutations along with rare *FLC* mutations explain a large fraction of flowering-time variation in *A. thaliana*. Here we document flowering time under different conditions in 20 *C. rubella* accessions from across the species' range. Similar to *A. thaliana*, vernalization, long photoperiods and elevated ambient temperature generally promote flowering. In this collection of *C. rubella* accessions, we did not find any obvious loss-of-function *FRI* alleles. Using mapping-by-sequencing with two strains that have contrasting flowering behaviors, we identified a splice-site mutation in *FLC* as the likely cause of early flowering in accession 1408. However, other similarly early *C. rubella* accessions did not share this mutation. We conclude that the genetic basis of flowering-time variation in *C. rubella* is complex, despite this very young species having undergone an extreme genetic bottleneck when it split from *C. grandiflora* a few tens of thousands of years ago.

FOR 2 decades, *Arabidopsis thaliana* has been the preeminent model for mechanistic studies of many aspects of plant development and physiology. In addition, it is extensively used to investigate the genetic basis of natural variation for adaptive traits such as the onset of flowering (Bergelson and Roux 2010; Koornneef and Meinke 2010; Weigel 2012). As in other species, flowering of *A. thaliana* is influenced by prolonged exposure to cold, which signals winter, by ambient temperature and long days, both of which indicate spring, as well as the age of the plant. An indication of the complexity of its regulation is that well over 100 genes affecting flowering time in the laboratory have been identified through mutant analysis (Srikanth and Schmid 2011).

*A. thaliana* is found throughout much of the Northern hemisphere, with a native range that extends from North Africa to the Arctic Circle, and from the Atlantic coast of Western Europe to Central Asia. Accordingly, *A. thaliana* accessions show a wide range of flowering-time behaviors when grown in climate chambers or in common gardens (Gazzani *et al.* 2003; Lempe *et al.* 2005; Werner *et al.* 2005; Brachi *et al.* 2010; Li *et al.* 2010; Hancock *et al.* 2011; Méndez-Vigo *et al.* 2011; Strange *et al.* 2011). In contrast to the multitude of laboratory-induced mutations that change flowering behavior, genetic analysis has identified only a handful of loci responsible for flowering variation among natural accessions. Chief among these are *FRIGIDA* (*FRI*) and *FLOWERING LOCUS C* (*FLC*), with *FRI* repressing flowering by allowing expression of *FLC*. Vernalization stably reduces the *FLC* transcription, thus reversing the action of *FRI*. Together, allelic differences at *FRI* and *FLC* can account for more than half of flowering-time variation when *A. thaliana* accessions are grown in constant, strongly flower-promoting conditions and without vernalization (Johanson


doi: 10.1534/genetics.112.143958
Manuscript received May 12, 2012; accepted for publication July 25, 2012
Available freely online through the author-supported open access option.
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.112.143958/-/DC1.
DNA sequences from this article have been deposited with the GenBank Library (JQ992992-JQ993051, JX003246-JX003248); short read data have been deposited with the European Nucleotide Archive (ENA) (ERP001651). Sequence variants have been deposited with Dryad (http://datadryad.org), doi:10.5061/dryad.4rq51.
[1]Present address: Department of Biotechnology, University of Delhi, Delhi-110007, India.
[2]Corresponding author: Department of Molecular Biology, Max Planck Institute for Developmental Biology, Spemannstr. 35, 72076 Tübingen, Germany. E-mail: weigel@weigelworld.org

*et al.* 2000; Le Corre *et al.* 2002; Gazzani *et al.* 2003; Michaels *et al.* 2003; Lempe *et al.* 2005; Shindo *et al.* 2005; Werner *et al.* 2005; Atwell *et al.* 2010; Brachi *et al.* 2010; Li *et al.* 2010; Méndez-Vigo *et al.* 2011; Salomé *et al.* 2011; Strange *et al.* 2011).

The ancestors of the two extant genera *Arabidopsis* and *Capsella* separated from each other about twice as long ago as *A. thaliana* split from other species in the genus *Arabidopsis* (Koch and Kiefer 2005). Reduced genetic diversity and the pattern of allele sharing suggest that *Capsella rubella* was founded about 30,000 to 50,000 years ago when a single *C. grandiflora* individual became self-compatible. The general picture is that most *C. rubella* alleles can be easily found in the *C. grandiflora* population and that loci typically have only one or two different major alleles (Foxe *et al.* 2009; Guo *et al.* 2009). Compared with *C. grandiflora*, *C. rubella* has a much wider distribution (Hurka and Neuffer 1997), which raises the question of how *C. rubella* could adapt so seemingly rapidly to a much wider range of environmental conditions than encountered by the genetically much more diverse *C. grandiflora*. In addition to new mutations, recombination could expose new epistatic interactions, allowing selection to act on standing variation (Barrett and Schluter 2008).

To address the extent and genetic basis of variation in a model adaptive trait, we have examined flowering-time variation in a set of *C. rubella* accessions. Using mapping-by-sequencing, we have identified a rare mutation in an *FLC* homolog as being responsible for rapid flowering in one particularly early accession. We demonstrate by transformation in *A. thaliana* that a splice site mutation greatly reduces the flowering-repressing activity of this allele.

## Materials and Methods

### *Plant material*

The *C. rubella* accessions studied (Table 1) have been described (Guo *et al.* 2009); they included the reference accessions MTE, for which a draft genome sequence has been produced by the U.S. Department of Energy Joint Genome Institute (DOE-JGI). The *A. thaliana* strain in the Col-0 background with the $FRI^{Sf2}$ and *flc*-3 alleles has been described (Lee *et al.* 1993; Michaels and Amasino 1999).

For flowering-time measurements, siblings from the same maternal family were used. Seeds were stratified for 7 days at 4° in 0.1% agar. For each treatment, we sowed six plants, most of which survived. Trays were moved to random positions in the growth rooms every 2 days to reduce positional effects. Flowering time was measured as total leaf number (TLN) and days to flowering (DTF).

### *Mapping-by-sequencing*

Early flowering $F_2$ plants from a cross of accession 1408 and the reference accession MTE were selected, and pooled genomic DNA was sequenced to 30-fold coverage on the Illumina GenomeAnalyzer IIx platform with 101-bp paired-end reads (59 million alignable reads, 11 Gb of sequence). SHORE (Ossowski *et al.* 2008) was used to discover single-nucleotide polymorphisms (SNPs) in the pool by alignment to a preliminary assembly of the MTE genome (CR_stitch_Feb15th.fasta) kindly provided by Jeremy Schmutz (DOE-JGI and HudsonAlpha), Stephen Wright and Khaled Hazzouri (University of Toronto), and Adrian Platts (McGill University). For annotation of variants, we used the *Capsella* v. 1.0 gene set released with Phytozome v8.0 (http://www.phytozome.net/capsella/). Using SHORE's scoring matrix approach optimized for heterozygous SNP detection and stringent filtering (requiring uniqueness of reads, minimum 10× coverage, minimum 20% allele frequency, SHORE SNP quality score >25), 354,080 SNPs could be identified (Supporting Information, Table S2). These SNPs were used as markers in SHOREmap (Schneeberger *et al.* 2009) to identify genomic regions with an excess of homozygous 1408 alleles. The final mapping interval was defined as having an allele frequency of at least 90% on both sides for 40 contiguous SNPs and fewer than 10 contiguous SNPs with allele frequency below 50% within this region.

### *Targeted DNA sequence analysis*

Individual young leaves were collected for DNA extraction using the cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987), followed by polymerase chain reaction (PCR) amplification with Pfu polymerase (Fermentas, St. Leon-Rot, Germany). Products from two independent reactions were mixed and directly Sanger sequenced on an ABI3730XL (Applied Biosystems, Foster City, CA). Primers for *FRI* were designed based on sequence of a *C. rubella* BAC (GenBank accession no. JX003248), and a 2.3-kb fragment covering the entire transcription unit was sequenced. Primers for *FLC* were designed based on *A. thaliana* genome sequence, and a 5.8-kb fragment encompassing most of the first exon through the 3′ UTR was sequenced. Sequences were assembled and inspected with Lasergene Seqman (DNASTAR, Madison, WI). See Table S1 for oligonucleotide primers.

Sequences were aligned using ClustalX v. 1.81 (Thompson *et al.* 1997), and alignments were refined manually. All polymorphic sites were individually confirmed again based on the original traces. DnaSP v. 4.10.9 (Rozas *et al.* 2003) was used to perform population analyses: levels of nucleotide diversity per site ($\pi$) (Nei 1987).

### *Expression analysis*

RNA was extracted from 2-week-old plants with the Plant RNeasy kit (Qiagen, Hilden, Germany) or with the TRIzol reagent (Invitrogen, Carlsbad, CA) and treated with RNase-free DNase I (Fermentas). Two to four micrograms of RNA was reverse transcribed using RevertAid first-strand cDNA Synthesis kit (Fermentas). PCRs were carried on in the presence of SYBR Green (Invitrogen, Carlsbad, CA) in 20-μl reactions. Amplification was monitored in real time with the Opticon continuous fluorescence detection system (MJ

**Table 1 Flowering times of *C. rubella* accessions**

| Accession | Country of origin | Latitude (°)[a] | Longitute (°)[a] | Elevation (m) | 16LD | | | | 16LDv | | | | 23LD | | | | 23SD | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | *n* | Range | Mean | Median | *n* | Range | Mean | Median | *n* | Range | Mean | Median | *n* | Range | Mean | Median |
| 690 | Esp | +36.15 | −5.58 | 50 | 6 | 65–74 | 71.8 | 74.0 | 6 | 32-35 | 34.5 | 35.0 | — | — | — | — | — | — | — | — |
| 697 | Ita | N/A | N/A | 950 | 6 | 56–56 | 56.0 | 56.0 | 6 | 35-35 | 35.0 | 35.0 | 6 | 36–40 | 38.7 | 39.0 | 6 | 63–71 | 68.7 | 69.5 |
| 698 | Ita | N/A | N/A | 1180 | 6 | 70–74 | 72.0 | 72.0 | 6 | 49-49 | 49.0 | 49.0 | — | — | — | — | — | — | — | — |
| 762 | Gre | +37.58 | +23.43 | N/A | 6 | 72–72 | 72.0 | 72.0 | 6 | 32-32 | 32.0 | 32.0 | — | — | — | — | — | — | — | — |
| 844 | Gre | N/A | N/A | 200 | 6 | 60–64 | 62.5 | 63.0 | 6 | 37-37 | 37.0 | 37.0 | — | — | — | — | — | — | — | — |
| 879 | Gre | +35.29 | +24.42 | 1200 | 6 | 51–51 | 51.0 | 51.0 | 6 | 32-32 | 32.0 | 32.0 | 6 | 28–29 | 28.3 | 28.0 | 6 | 63–77 | 69.8 | 69.5 |
| 907 | Gre | +39.40 | +19.48 | 40 | 6 | 61–61 | 61.0 | 61.0 | 5 | 51-51 | 51.0 | 51.0 | — | — | — | — | — | — | — | — |
| 925 | Gre | +39.40 | +20.51 | 450 | 6 | 98–100 | 98.3 | 98.0 | 6 | 51-51 | 51.0 | 51.0 | — | — | — | — | — | — | — | — |
| 1204 | Esp | +28.19 | −16.34 | 1000 | 6 | 72–74 | 72.3 | 72.0 | 6 | 37-37 | 37.0 | 37.0 | — | — | — | — | — | — | — | — |
| 1207 | Esp | +28.19 | −16.34 | 1100 | 6 | 56–58 | 57.7 | 58.0 | 6 | 29-29 | 29.0 | 29.0 | 6 | 70–85 | 77.3 | 78.0 | 6 | 58–103 | 75.2 | 70.5 |
| 1208 | Esp | +28.19 | −16.34 | 100 | 6 | 51–56 | 53.5 | 53.5 | 6 | 29-29 | 29.0 | 29.0 | 4 | 70–95 | 78.5 | 74.5 | 6 | 68–106 | 88.7 | 87.5 |
| 1209 | Esp | +28.19 | −16.34 | 1200 | 6 | 70–72 | 70.3 | 70.0 | 6 | 32-32 | 32.0 | 32.0 | — | — | — | — | — | — | — | — |
| 1215 | Esp | +28.19 | −16.19 | 0 | 6 | 60–65 | 62.3 | 63.0 | 6 | 27-29 | 28.7 | 29.0 | 6 | 44–67 | 56.2 | 59.5 | 1 | 103 | 103 | 103 |
| 1311 | Fra | +42.53 | −0.06 | N/A | 6 | 70–81 | 76.5 | 77.0 | 6 | 27-29 | 28.0 | 28.0 | — | — | — | — | — | — | — | — |
| 1377 | Arg | −34.40 | −58.30 | 10 | 6 | 65–70 | 66.7 | 65.0 | 6 | 32-32 | 32.0 | 32.0 | 5 | 36–72 | 48.6 | 36.0 | 5 | 52–91 | 70.0 | 68.0 |
| 1408 | Gre | +35.29 | +24.42 | 170 | 6 | 51–52 | 51.3 | 51.0 | 6 | 29-29 | 29.0 | 29.0 | 4 | 65–72 | 68.0 | 67.5 | 4 | 56–56 | 56.0 | 56.0 |
| 1453 | Ita | +43.28 | +11.02 | N/A | 6 | 74–77 | 75.5 | 75.5 | 6 | 42-44 | 43.0 | 43.0 | 6 | 60–79 | 71.5 | 79.0 | 1 | 77 | 77.0 | 77.0 |
| 1482 | Aus | −31.56 | +115.50 | N/A | 6 | 77–79 | 77.7 | 77.0 | 6 | 32-35 | 33.0 | 32.0 | — | — | — | — | — | — | — | — |
| MTE | Ita | N/A | N/A | N/A | 6 | 100–106 | 102.5 | 103.0 | 6 | 42-44 | 43.3 | 44.0 | — | — | — | — | — | — | — | — |
| RIAH | Ita | N/A | N/A | N/A | 6 | 70–74 | 71.3 | 71.0 | 6 | 32-32 | 32.0 | 32.0 | — | — | — | — | — | — | — | — |

N/A, information not available.

[a] + or − indicates Northern and Southern latitudes, and Eastern and Western longitudes.

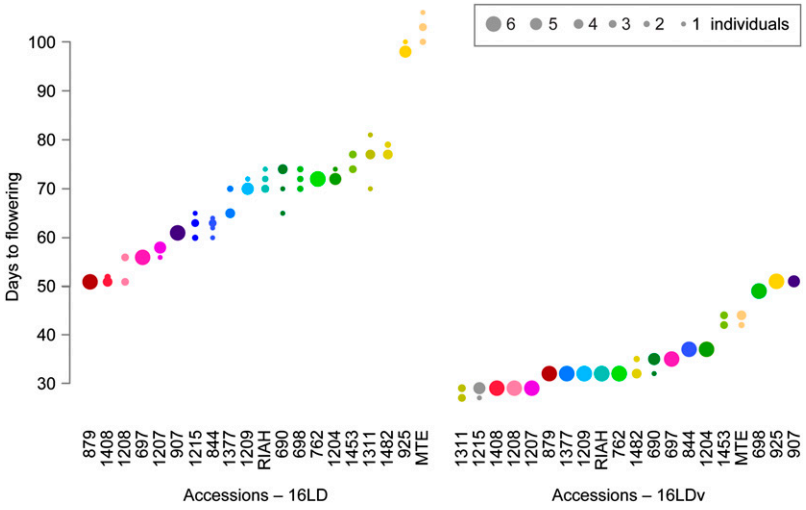


**Figure 1** Distribution of flowering time among 20 *C. rubella* accessions in 16LD and16LDv. Note the differences in scale. Accessions, which are ordered by mean flowering time in each condition, are color coded for comparison. For each accession, the number of individuals that flowered on a given day is indicated.

Research, Reno, NV). Threshold cycles (cT) were based on a reaction reaching a specific fluorescence intensity in the log-linear phase of the amplification curve.

For the analysis of splice forms, a fragment of about 200 bp was amplified by PCR from pooled complementary DNA (cDNA) obtained from four primary transformants for each transgenic line and digested for 2 hr with 1 μl of Fast Digest *Bst*XI (Fermentas). See Table S1 for oligonucleotide primers.

### *Transgenic lines*

*FLC* genomic fragments covering the entire open reading frame and introns were amplified using Phusion polymerase (Thermo Fisher Scientific, Waltham, MA). These fragments were placed under the control of the constitutive CaMV 35S promoter in the pGREEN-derived (Hellens *et al.* 2000) vector pFK210 and introduced into different *A. thaliana* genotypes by *Agrobacterium tumefaciens*-mediated transformation (Weigel and Glazebrook 2002).

### *Statistical analyses*

Mean, median, and standard deviation were calculated using Microsoft Excel (Microsoft, Redmond, WA). Differences between groups in Figure 7B were tested by ANOVA using Tukey's test, as implemented in Excel. Spearman's rank-order correlation coefficient $r_s$ was determined in R (http://www.r-project.org/).

## Results

### *Flowering-time variation in C. rubella*

We analyzed the response of *C. rubella* accessions to long-day (16 hr light) and short-day (8 hr light) photoperiods and to vernalization for 6 weeks at 4° in short-day photoperiods, after seeds had been germinated on soil in 16°. Twenty accessions were grown in 16° long days without vernalization (16LD) and in 16° long days with vernalization (16LDv) (Table 1). A subset of eight accessions was grown in two 23° environments with contrasting photoperiods, long days (23LD), and short days (23SD) (Table 1). As in *A. thaliana* (Koornneef *et al.* 1991; Gazzani *et al.* 2003; Lempe *et al.* 2005), DTF and TLN were correlated (16LD, $r_s = 0.62$, $P = 0.003$; 16LDv, $r_s = 0.80$, $P << 0.001$). Below, we report only DTF data.

Accessions flowered on average much earlier after vernalization (Figure 1). In 16LD, onset of flowering ranged from 51 to 103 days (mean, 69; median, 71). In 16 LDv, it ranged from 28 to 51 days (mean, 36; median, 33). While

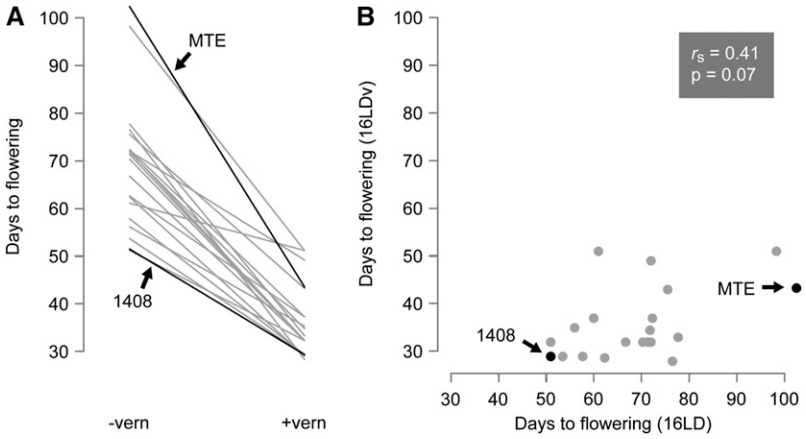


**Figure 2** Relationship of flowering-time variation and vernalization response. (A) Reaction norms of flowering time in 16LD without and with vernalization (−vern, +vern). (B) Correlation of mean flowering times with and without vernalization; same data as in Figure 2A.

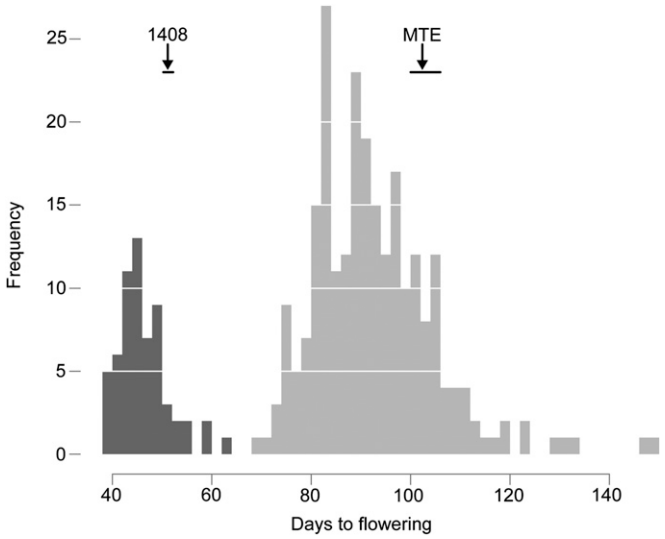


**Figure 3** Flowering time in 16LD among 305 $F_2$ individuals from the MTE × 1408 cross. Average and range of flowering times of grandparents are indicated. Darker shading indicates the 61 individuals used for bulked segregant mapping.

all accessions responded to vernalization, there was considerable variation in the extent of the response, as seen most clearly in a reaction norm representation (Figure 2A). Late-flowering accessions tended to respond more strongly to vernalization than early flowering ones, but DTF variation in 16LD explained about 17% of DTF variation in 16LDv (Figure 2B). In addition to vernalization, both longer photoperiods and increased ambient temperature tended to accelerate flowering (Table 1).

### *FLC underlies a major effect locus for flowering time*

The reference accession MTE, from Italy, was the last accession to flower in 16LD, after 103 days on average, which was twice as late as accession 1408, from Greece (Table 1). Following vernalization in 16LDv, MTE was still late, flowering after 43 days, but it was no longer the last accession to flower. A robust vernalization response was also seen in accession 1408, which flowered in 16LDv after 29 days, as did several other accessions.

To begin to investigate the genetic basis of flowering-time variation in *C. rubella*, we crossed the late-flowering accession MTE with the early flowering accession 1408. In an $F_2$ population, about one-quarter of plants flowered early, similar to 1408, and three- quarters flowered late, similar to MTE (Figure 3), suggesting that a major, recessively acting gene controlled early flowering of accession 1408.

MTE is the reference accession for the ongoing *C. rubella* genome sequencing effort (collaboration with Jeremy Schmutz from the Joint Genome Institute of the U.S. Department of Energy and the HudsonAlpha Institute for Biotechnology, Stephen Wright and Khaled Hazzouri of the University of Toronto, and Adrian Platts of McGill University). The MTE draft genome sequence in turn enables the discovery of polymorphisms between MTE and 1408, including polymorphisms that are potentially causal for the flowering-time differences between the two accessions. With this rationale in mind, we selected from 305 $F_2$ plants of the MTE × 1408 cross 61 individuals that had flowered after 65 days in 16 LD (Figure 3 and Table S3). We extracted a single pool of genomic DNA from these plants and sequenced the pooled DNA on the Illumina GAIIx platform.

Sequencing reads were aligned to a preliminary genome assembly of *C. rubella* MTE. The proportion of sequence reads that could not be mapped on the MTE genome was 7.3%, which is similar to what we have found with *A. thaliana* accessions (Cao *et al.* 2011). As described (Schneeberger *et al.* 2009), polymorphisms that can be used for genetic mapping, including SNPs, can be identified directly from $F_2$ data, if one allows for both homo- and heterozygous SNPs. Accordingly, SNPs characteristic for 1408 were discovered from the bulked segregant data. These SNPs were then used to identify regions enriched for either 1408 or MTE SNPs (Schneeberger *et al.* 2009). Rearrangements and indels were not considered in this step. With this strategy, we identified on chromosome 6 of the genome assembly a region of 333 kb that contained over 600 positions where over 90% of reads supported a nonreference polymorphism (Figure 4 and Table S4). The overall density of 2.0 SNPs/kb was close to the genome-wide average of 2.3 SNPs/kb.

Using all reads from the bulked segregant pools that mapped to this region, we annotated 1408 variants relative to the *Capsella* v. 1.0 gene set (http://www.phytozome.net/capsella/), which contained 90 protein-coding genes in the final mapping interval. The algorithms implemented in SHORE v. 0.8 (Ossowski *et al.* 2008) and pindel v. 0.2.4s (Ye *et al.* 2009) detected 652 SNPs, 184 small indels of length 1 to 3 bp, 25 longer structural variants (SVs) of length 7 to 679 bp, and 109 highly variable regions, recognizable as simultaneous insertions and deletions (Table S4). Copy number variants (CNVs) and inversions were not detected. Fifty-one SNPs were predicted to cause nonsynonymous amino acid substitutions in 35 genes, with one SNP altering a stop codon. Five SVs affected the coding regions of four genes, and 14 highly variable regions affected the coding regions of 11 genes. Close to the middle of the final mapping interval from position 2.968 to 3.301 Mb, at position 3.112 to 3.120 Mb, was the only *C. rubella* homolog of *A. thaliana FLC*, a well-known flowering repressor (Sheldon *et al.* 1999; Michaels *et al.* 2003). None of the other genes in this interval were closely related to homologs of known flowering-time genes.

In *A. thaliana*, there is large variation in *FLC* levels, and accessions with weak *FLC* alleles often have very little *FLC* expression (Lempe *et al.* 2005; Shindo *et al.* 2005). Variation in *FLC* expression was more moderate in *C. rubella* accessions. MTE had the highest *FLC* levels, consistent with it being the latest accession, but accession 1408 did not have the least *FLC* expression (Figure 5A). Across all accessions, there was no clear correlation of *FLC* expression and flowering in unvernalized plants (Figure 5A), but a modest

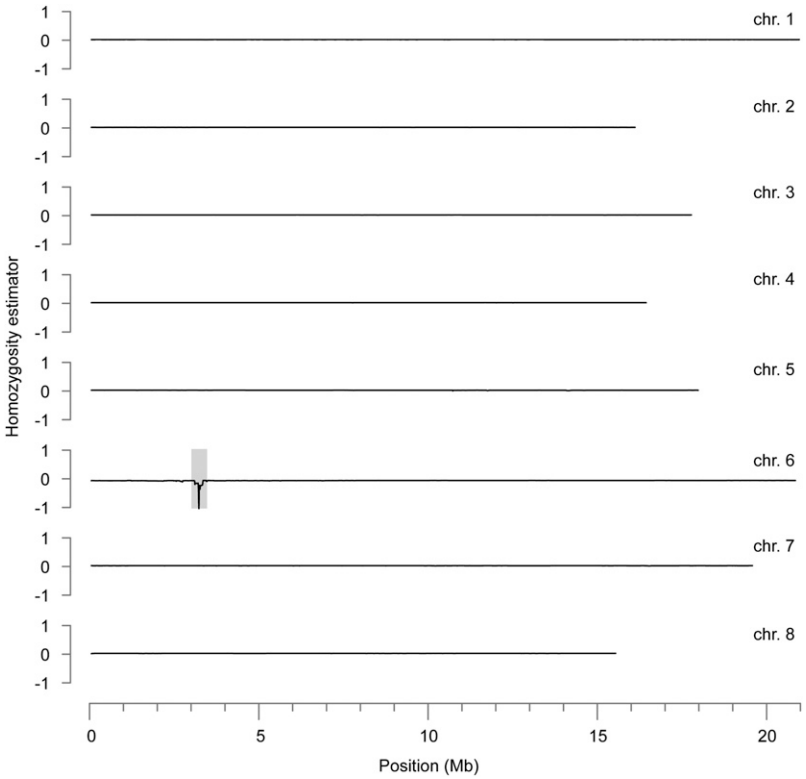


**Figure 4** SHOREmap analysis of early flowering QTL. The homozygosity estimator is 0 at even allele frequencies of both parents, 1 when homozygous for late-flowering accession MTE, and −1 when homozygous for early flowering accession 1408 (Schneeberger *et al.* 2009). Sliding windows of 100 kb with step size 10 kb were used. The region of chromosome 6 enriched at over 600 markers indicative of 1408 is indicated in gray. The eight largest scaffolds of a preliminary *C. rubella* genome assembly are shown, corresponding to the majority of the eight chromosomes.

correlation of *FLC* and vernalization response (Figure 5B). As expected, expression of *FLC* was reduced after vernalization, although to quite a different extent in different accessions, which is clearly seen in a reaction norm representation (Figure 5C).

Because it was unclear whether differences of *FLC* expression were likely to be the cause for earlier flowering of accession 1408 compared to MTE, we compared the *FLC* DNA sequence of the two accessions. There is an ACAG-to-AGAG substitution at the very end of intron 6 that shifts the splice acceptor site so that two bases are inserted into the *FLC* mRNA of accession 1408, as revealed by sequencing of cDNA. This leads to a frame shift and truncation of the open reading frame, removing the last 35 of 198 amino acids (Figure 6).

To test whether this polymorphism could indeed be responsible for the flowering-time difference between MTE and 1408, we introduced both *FLC* alleles into *A. thaliana* $FRI^{Sf\text{-}2}$ *flc*-3 plants. Because we did not know the extent of regulatory sequences in *C. rubella*, we expressed the *FLC* alleles under the control of the constitutive cauliflower mosaic virus 35S promoter. The MTE allele strongly delayed flowering, as did overexpression of a wild-type *FLC* copy from *A. thaliana*, and both were similarly effective in repressing the flowering activators *FT* and *SOC1* (Figure 7, A and B). In contrast, the 1408 allele delayed flowering by only

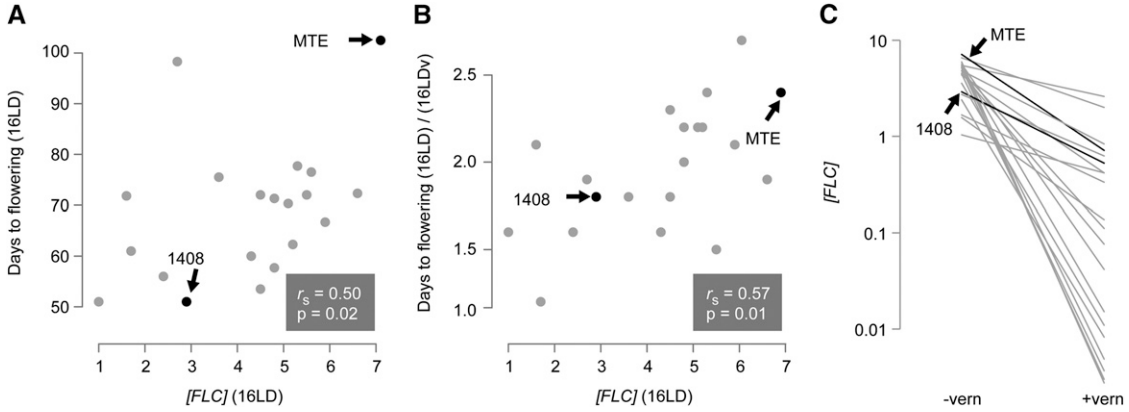


**Figure 5** Relationship between *FLC* expression measured by qRT–PCR and flowering time of accessions. (A) Correlation between *FLC* expression and flowering without vernalization. The lowest value was arbitrarily set to 1. (B) Correlation between *FLC* expression and vernalization response. (C) Effect of vernalization (vern) on *FLC* expression. The difference between *FLC* and *BETA-TUBULIN*, ΔcT, was used to calculate *FLC* expression, assuming PCR efficiency was 100% for both genes. Two biological replicates, each with two technical replicates, were analyzed.

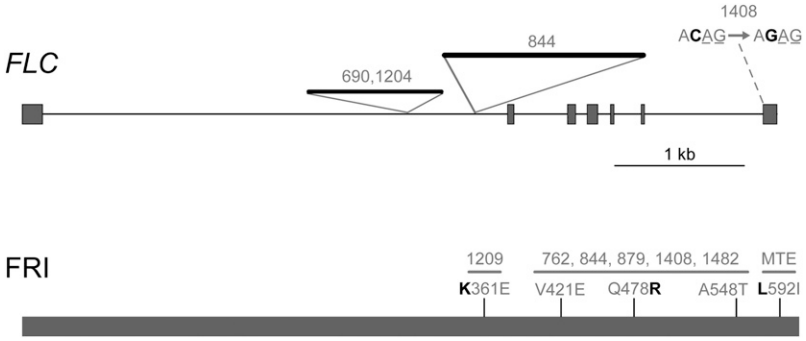


**Figure 6** Sequence variation of *FLC* and *FRI* in *C. rubella*. In addition to the *FLC* point mutation at the end of the last intron in accession 1408, two insertions that are segregating in the population, are shown. Thin lines indicate introns; thick lines protein-coding sequences. The C → G substitution creates a new splice acceptor site. PCR amplification and sequencing of cDNA confirmed that 2 bases are inserted into the *FLC* mRNA in accession 1408. There was no apparent heterogeneity in the sequence, suggesting that the canonical splice form, if it exists at all, is rare. For FRI, amino acid residues that appear to be ancestral based in several Brassicaceae (Irwin *et al.* 2012) are shown in boldface type; numbers indicate position in the peptide sequence. Shaded numbers indicate affected accessions.

a few days, from 25.2 to 29.6 days, and it had a much smaller effect on *FT* and *SOC1* expression (Figure 7, A and B). *FLC* transcript levels were similar in all three transgenic lines (Figure 7B). An analysis of cDNA obtained from plants expressing either the MTE or 1408 allele showed that the 1408 allele produced only the misspliced version of the *FLC* transcript (Figure 7C).

### *FLC and variation in flowering-time and vernalization response*

Several other accessions flowered as early as 1408 both with and without vernalization (Table 1 and Figure 2). A survey of *FLC* sequences, however, did not reveal any additional accessions with the 1408 mutation. Compared to *C. grandiflora*, there were very few *FLC* polymorphisms in *C. rubella*. All *FLC* sequences from *C. rubella* clustered together and formed a group that was distinct from *C. grandiflora FLC* alleles (Guo *et al.* 2009). This strongly suggested that the 1408 mutation arose very recently, after the split of *C. rubella* and *C. grandiflora*.

In *A. thaliana*, large insertions into the first intron are a common cause for reduced *FLC* activity (Gazzani *et al.* 2003; Michaels *et al.* 2003; Liu *et al.* 2004; Lempe *et al.* 2005; Shindo *et al.* 2005; Strange *et al.* 2011; Sánchez-Bermejo *et al.* 2012). Three *C. rubella* accessions had over 1-kb-long insertions into the first intron, but none of them stood out as being particularly early flowering or having a particularly weak vernalization response (Table 1 and Figure 6). We also surveyed the accessions for changes in the open reading frame of *FRI*, which is located in a region syntenic to chromosome 5 of *A. thaliana* (Figure S1) (Irwin *et al.* 2012). The open reading frames were intact in all 20 accessions. However, there were several nonsynonymous substitutions (Figure 6), and $\pi$ was much higher for *FRI* than for *FLC* (0.004 compared with 0.0004).

## Discussion

### *Flowering-time variation in the Brassicaceae*

Apart from disease resistance, flowering time is perhaps the most comprehensively studied adaptive trait in *A. thaliana*. In both laboratory and field conditions, there is a wide range in the onset of flowering. Much of this variation can be

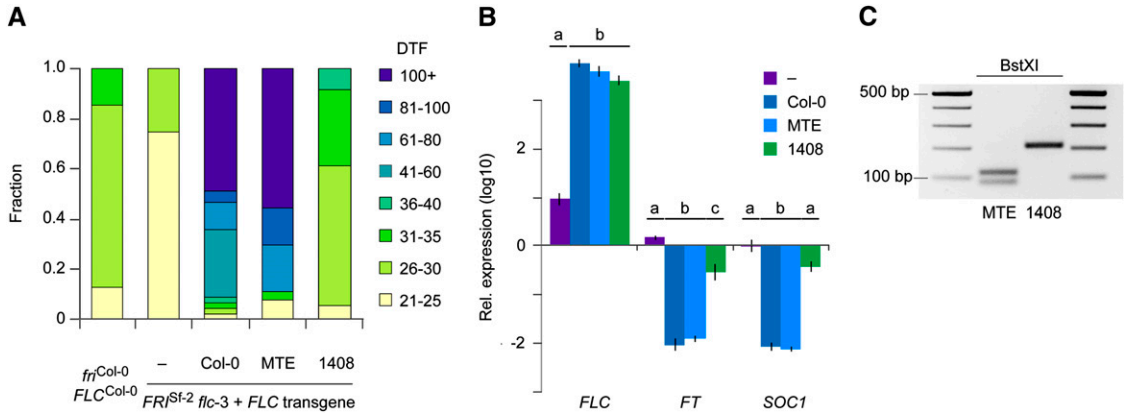


**Figure 7** Functional analysis of *C. rubella FLC* alleles in *A. thaliana*. (A) Flowering time of plants grown in 23LD. The three left strains are controls. $fri^{Col\text{-}0}$ $FLC^{Col\text{-}0}$ is the Col-0 reference strain, which carries a functional *FLC* allele, but a naturally inactive *FRI* allele (Johanson *et al.* 2000). $FRI^{Sf\text{-}2}$ *flc*-3 carries an introgression of the functional *FRI* allele from the Sf-2 accession and has also an induced mutation at the *FLC* locus (Michaels and Amasino 1999). The three right strains are transgenic lines, expressing the indicated *FLC* allele from the constitutive CaMV 35S promoter. The effect of the *C. rubella* MTE allele is similar to that of the fully functional *A. thaliana* Col-0 allele. $n \geq 30$. (B) Expression of *FLC* and of two downstream flowering regulators in transgenic and non-transgenic $FRI^{Sf\text{-}2}$ *flc*-3 plants, determined by qRT–PCR using the $2^{-\Delta\Delta cT}$ method. Expression levels were normalized to those of wild-type Col-0 plants. Averages for four individuals for controls and eight for transgenic lines are shown. Error bars indicate standard errors of the mean. Differences between groups are significant at $P < 0.05$. (C) cDNA analysis of MTE and 1408 alleles expressed in transgenic *A. thaliana* plants. The splicing variant generated by the SNP in 1408 abolishes a *Bst*XI restriction site in the amplified cDNA fragment.

explained by allelic variation at *FRI* along with *FLC* and its homologs. Consistent with the role of *FRI* and *FLC* in repressing flowering in unvernalized plants, much of the flowering variation disappears after vernalization (recently reviewed by Weigel 2012).

Because allelic variation at *FRI* and *FLC* is so common in *A. thaliana*, their homologs have been prime candidates for loci affecting flowering time in other Brassicaceae (Axelsson *et al.* 2001; Tadege *et al.* 2001; Lou *et al.* 2007; Zhao *et al.* 2010a; Uptmoor *et al.* 2012). There is evidence for *FRI* or *FLC* homologs being responsible for quantitative trait loci (QTL) affecting flowering in *A. lyrata*, *Brassica napus*, and *B. rapa*, with the situation being less clear for *B. oleracea* (Schranz *et al.* 2002; Pires *et al.* 2004; Long *et al.* 2007; Okazaki *et al.* 2007; Razi *et al.* 2007; Kuittinen *et al.* 2008; Zhao *et al.* 2010b; Wang *et al.* 2011). While the potential molecular causes for a change in *FLC* function in these cases remain uncertain, splice-site mutations in *FLC* homologs have been linked to earlier flowering in *B. rapa* and *Capsella bursa-pastoris* (Slotte *et al.* 2009; Yuan *et al.* 2009; Hu *et al.* 2011). However, in all these cases there has been little direct proof for a causal role of *FLC* or *FRI* alleles in flowering-time variation. An induced mutation in the *FLC* homolog *PERPETUAL FLOWERING 1* (*PEP1*) changes the perennial flowering habit of *Arabis alpina*. Its response to vernalization is different from the one of *A. thaliana FLC*, pointing to diversification in the function of *FLC* homologs in the Brassicaceae (Wang *et al.* 2009).

### *Flowering-time variation in C. rubella*

*Capsella* is more distantly related to *A. thaliana* than the other *Arabidopsis* species, but more closely than *Brassica* and *Arabis*. *Capsella* is a small genus with only five species, including one tetraploid species, *C. bursa-pastoris*, one of the most frequent and widespread cosmopolitan weeds (Hurka and Neuffer 1997; Ceplitis *et al.* 2005; Hurka *et al.* 2012). The four diploid species include the self-compatible *C. rubella*, which originated only a few tens of thousands of years ago from the self-incompatible *C. grandiflora* (Foxe *et al.* 2009; Guo *et al.* 2009). *C. grandiflora* has a restricted distribution, in Western Greece and Albania and, rarely, Northern Italy, while *C. rubella* has spread across the Mediterranean, and followed European settlers into the New World and Australia (Hurka and Neuffer 1997). In this work, we have documented flowering-time variation in *C. rubella* under laboratory conditions. Similar to *A. thaliana*, flowering of many *C. rubella* accessions is accelerated by vernalization, longer photoperiods, and increased ambient temperature. As in *A. thaliana*, later-flowering *C. rubella* accessions tend to respond more strongly to vernalization than early flowering ones (Lempe *et al.* 2005; Shindo *et al.* 2005). However, the correlation between flowering-time and vernalization response, or between flowering-time and *FLC* expression, is less pronounced in *C. rubella*. This could be due to biased sampling, or it might reflect functional differences between *C. rubella* and *A. thaliana*. Given the different ranges of the two species, the second hypothesis is well worth further testing.

*FRI* alleles with premature stop codons explain a large fraction of flowering-time variation in *A. thaliana* (Le Corre *et al.* 2002; Lempe *et al.* 2005; Shindo *et al.* 2005; Werner *et al.* 2005; Atwell *et al.* 2010) In *C. rubella*, we did not find any potential loss-of-function allele at *FRI*, even though the *C. rubella* accessions surveyed flower faster than many *A. thaliana* accessions. An indication for the functionality of *FRI* is that all *C. rubella* accessions analyzed responded to vernalization. In *A. thaliana*, structural variation at *FLC* itself appears to be often responsible for differences in the magnitude of the vernalization response (Michaels *et al.* 2003; Lempe *et al.* 2005; Shindo *et al.* 2006; Sánchez-Bermejo *et al.* 2012). In *C. rubella*, we did not observe an obvious correlation between structural variation at *FLC* and vernalization response.

### *Flowering-time variation through independent FLC mutations*

The reference accession MTE turned out to be one of the latest-flowering *C. rubella* accessions we surveyed. The segregation pattern in the $F_2$ of a cross to one of the earliest flowering accession in our collection, 1408, suggested the presence of a single major effect QTL responsible for much of the flowering-time difference between the two accessions. To rapidly fine map this QTL, we exploited recent advances in mapping-by-sequencing methods that do not require separate sequence analysis of the parental genomes (Schneeberger *et al.* 2009). We found that the gene most likely to be responsible for the QTL is the *C. rubella* homolog of *FLC*. While we cannot exclude that other genes in the interval contribute to this QTL, none of the 89 other genes in the mapping interval is closely related to a known flowering-time regulator. Moreover, we note that it is difficult to draw firm conclusions about the quantitative differences in the activity of the MTE and 1408 alleles from the *A. thaliana* experiments, since there might be additional *trans*- and *cis*-factors that modify *FLC* activity between the two species.

The likely causal polymorphism in *FLC* affects splicing, which in turn removes the C-terminal 35 amino acids. That most of the open reading frame remains intact may explain why accession 1408 retains a robust vernalization response. In *A. thaliana*, reduction- or loss-of-function alleles of *FLC* either have insertions in the first, large intron, or they carry premature stop codons that can lead to alternative splicing (Michaels *et al.* 2003; Liu *et al.* 2004; Lempe *et al.* 2005; Shindo *et al.* 2005; Werner *et al.* 2005; Méndez-Vigo *et al.* 2011; Sánchez-Bermejo *et al.* 2012). However, alleles with splice-site mutations have, to our knowledge, not been described in *A. thaliana*. It is striking that *FLC* alleles with splice-site mutations have now been found in three other species, in *B. rapa*, *C. bursa-pastoris*, and *C. rubella* (this work; Slotte *et al.* 2009; Yuan *et al.* 2009; Hu *et al.* 2011).

Although several other accessions show a similar flowering behavior as in 1408, the 1408 *FLC* allele is not shared by

any other accession. Moreover, the pattern of polymorphisms in *FLC* alleles suggests that the 1408 mutation arose only after the split from *C. grandiflora* (Figure S2). Because of the strong recent bottleneck experienced by *C. rubella* (Foxe *et al.* 2009; Guo *et al.* 2009), there is only limited standing variation in this species, and new mutations likely play an important role in the adaptation to different environments. There is precedence for *FLC*, and to a lesser extent, *FRI*, loss-of-function alleles being rare. Although few individual loss-of-function alleles segregate at appreciable frequency in *A. thaliana* flowering, collectively, such alleles are quite frequent. There is good evidence that the more common *FRI* loss-of-function alleles have increased in frequency due to selection for early flowering (Toomajian *et al.* 2006), a conclusion that is supported by experimental selection experiments (Scarcelli and Kover 2009). In animals, most cases of parallel evolution are from closely related species, but there are also examples of the same gene being responsible for genetic variation in more distantly related taxa, the most notorious one being the melanocortin-1 receptor gene (*Mc1r*), which underlies changes in pigmentation in reptiles, birds, and mammals (Wood *et al.* 2005; Hoekstra 2006; Arendt and Reznick 2008; Protas and Patel 2008; Gompel and Prud'homme 2009; Elmer and Meyer 2011). Similarly, there are several cases of parallel evolution in flower pigmentation (Quattrocchio *et al.* 1999; Schwinn *et al.* 2006; Whittall *et al.* 2006; Hoballah *et al.* 2007; Streisfeld and Rausher 2009; Des Marais and Rausher 2010; Smith and Rausher 2011). How broadly *FLC* contributes to flowering-time variation in *C. rubella* is, however, not known yet. Thus, further studies are needed to determine whether it is appropriate to speak of parallel evolution when comparing the basis of flowering-time variation in *A. thaliana* and *C. rubella*.

In summary, we have found substantial flowering-time variation in *C. rubella*, although its extent is smaller than in *A. thaliana*, perhaps reflecting the more restricted geographic range of *C. rubella* (Hurka and Neuffer 1997; Hoffmann 2002). In one accession, a major-effect QTL affecting flowering maps *FLC*. In experimental crosses of *A. thaliana*, a small number of major effect QTL, including *FRI*, *FLC*, the *FLC* homolog *FLOWERING LOCUS M* (*FLM*)/*MADS AFFECTING FLOWERING 1* (*MAF1*), the *MAF2-5* cluster, and *FLOWERING LOCUS T* (*FT*), explain much of the variation in flowering time (Brachi *et al.* 2010; Salomé *et al.* 2011; Strange *et al.* 2011). Given the reduced genetic diversity in *C. rubella* (Foxe *et al.* 2009; Guo *et al.* 2009), it will be interesting to determine how many additional loci are responsible for flowering-time variation.

## Acknowledgments

We thank Jeremy Schmutz (DOE-JGI and HudsonAlpha), Stephen Wright and Khaled Hazzouri (University of Toronto), and Adrian Platts (McGill University) for access to a preliminary version of the *C. rubella* genome assembly. We thank Sureshkumar Balasubramanian, Patrice Salomé, Jia-Wei Wang, Xuan Zhao, and Zhong Zhao for advice, Dan Koenig and Patrice Salomé for comments on the manuscript, Korbinian Schneeberger for help with SHOREmap analyses, and Christa Lanz for sequencing support. This work was supported by Bundesministerium für Bildung und Forschung (BMBF) grant Next Generation Forward Genetics for Crop Plants (NUGGET), a Gottfried Wilhelm Leibniz Award to the Deutsche Forschungsgemeinschaft (DFG), and the Max Planck Society.

## Literature Cited

Arendt, J., and D. Reznick, 2008 Convergence and parallelism reconsidered: What have we learned about the genetics of adaptation? Trends Ecol. Evol. 23: 26–32.

Atwell, S., Y. S. Huang, B. J. Vilhjalmsson, G. Willems, M. Horton *et al.*, 2010 Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature 465: 627–631.

Axelsson, T., O. Shavorskaya, and U. Lagercrantz, 2001 Multiple flowering time QTLs within several *Brassica* species could be the result of duplicated copies of one ancestral gene. Genome 44: 856–864.

Barrett, R. D., and D. Schluter, 2008 Adaptation from standing genetic variation. Trends Ecol. Evol. 23: 38–44.

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

Brachi, B., N. Faure, M. Horton, E. Flahauw, A. Vazquez *et al.*, 2010 Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genet. 6: e1000940.

Cao, J., K. Schneeberger, S. Ossowski, T. Gunther, S. Bender *et al.*, 2011 Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nat. Genet. 43: 956–963.

Ceplitis, A., Y. Su, and M. Lascoux, 2005 Bayesian inference of evolutionary history from chloroplast microsatellites in the cosmopolitan weed *Capsella bursa-pastoris* (Brassicaceae). Mol. Ecol. 14: 4221–4233.

Des Marais, D. L., and M. D. Rausher, 2010 Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in *Ipomoea*. Evolution 64: 2044–2054.

Doyle, J. J., and J. L. Doyle, 1987 A rapid DNA isolation procedure from small quantities of fresh leaf tissues. Phytochem. Bull. 19: 11–15.

Elmer, K. R., and A. Meyer, 2011 Adaptation in the age of ecological genomics: insights from parallelism and convergence. Trends Ecol. Evol. 26: 298–306.

Foxe, J. P., T. Slotte, E. A. Stahl, B. Neuffer, H. Hurka *et al.*, 2009 Recent speciation associated with the evolution of selfing in *Capsella*. Proc. Natl. Acad. Sci. USA 106: 5241–5245.

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

Gompel, N., and B. Prud'homme, 2009 The causes of repeated genetic evolution. Dev. Biol. 332: 36–47.

Guo, Y. L., J. S. Bechsgaard, T. Slotte, B. Neuffer, M. Lascoux *et al.*, 2009 Recent speciation of *Capsella rubella* from *Capsella grandiflora*, associated with loss of self-incompatibility and an extreme bottleneck. Proc. Natl. Acad. Sci. USA 106: 5246–5251.

Hancock, A. M., B. Brachi, N. Faure, M. W. Horton, L. B. Jarymowycz *et al.*, 2011 Adaptation to climate across the *Arabidopsis thaliana* genome. Science 334: 83–86.

Hellens, R. P., E. A. Edwards, N. R. Leyland, S. Bean, and P. M. Mullineaux, 2000 pGreen: a versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. Plant Mol. Biol. 42: 819–832.

Hoballah, M. E., T. Gubitz, J. Stuurman, L. Broger, M. Barone *et al.*, 2007 Single gene-mediated shift in pollinator attraction in *Petunia*. Plant Cell 19: 779–790.

Hoekstra, H. E., 2006 Genetics, development and evolution of adaptive pigmentation in vertebrates. Heredity 97: 222–234.

Hoffmann, M. H., 2002 Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). J. Biogeogr. 29: 125–134.

Hu, G. L., Z. L. Hu, Y. Li, F. Gu, Z. P. Zhao *et al.*, 2011 A splicing site mutation in *BrpFLC1* and repressed expression of *BrpFLC* genes are associated with the early flowering of purple flowering stalk. Russ. J. Plant Physiol. 58: 431–438.

Hurka, H., and B. Neuffer, 1997 Evolutionary processes in the genus *Capsella* (Brassicaceae). Plant Syst. Evol. 206: 295–316.

Hurka, H., N. Friesen, D. A. German, A. Franzke, and B. Neuffer, 2012 'Missing link' species *Capsella orientalis* and *Capsella thracica* elucidate evolution of model plant genus Capsella (Brassicaceae). Mol. Ecol. 21: 1223–1238.

Irwin, J. A., C. Lister, E. Soumpourou, Y. Zhang, E. C. Howell *et al.*, 2012 Functional alleles of the flowering time regulator *FRIGIDA* in the *Brassica oleracea* genome. BMC Plant Biol. 12: 21.

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

Koch, M. A., and M. Kiefer, 2005 Genome evolution among cruciferous plants: a lecture from the comparison of the genetic maps of three diploid species—*Capsella rubella*, *Arabidopsis lyrata* subsp. *petraea*, and *A. thaliana*. Am. J. Bot. 92: 761–767.

Koornneef, M., and D. Meinke, 2010 The development of *Arabidopsis* as a model plant. Plant J. 61: 909–921.

Koornneef, M., C. J. Hanhart, and J. H. van der Veen, 1991 A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. Mol. Gen. Genet. 229: 57–66.

Kuittinen, H., A. Niittyvuopio, P. Rinne, and O. Savolainen, 2008 Natural variation in *Arabidopsis lyrata* vernalization requirement conferred by a *FRIGIDA* indel polymorphism. Mol. Biol. Evol. 25: 319–329.

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. Mol. Biol. Evol. 19: 1261–1271.

Lee, I., A. Bleecker, and R. Amasino, 1993 Analysis of naturally occurring late flowering in *Arabidopsis thaliana*. Mol. Gen. Genet. 237: 171–176.

Lempe, J., S. Balasubramanian, S. Sureshkumar, A. Singh, M. Schmid *et al.*, 2005 Diversity of flowering responses in wild *Arabidopsis thaliana* strains. PLoS Genet. 1: 109–118.

Li, Y., Y. Huang, J. Bergelson, M. Nordborg, and J. O. Borevitz, 2010 Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA 107: 21199–21204.

Liu, J., Y. He, R. Amasino, and X. Chen, 2004 siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in *Arabidopsis*. Genes Dev. 18: 2873–2878.

Long, Y., J. Shi, D. Qiu, R. Li, C. Zhang *et al.*, 2007 Flowering time quantitative trait loci analysis of oilseed brassica in multiple environments and genomewide alignment with Arabidopsis. Genetics 177: 2433–2444.

Lou, P., J. Zhao, J. S. Kim, S. Shen, D. P. Del Carpio *et al.*, 2007 Quantitative trait loci for flowering time and morphological traits in multiple populations of *Brassica rapa*. J. Exp. Bot. 58: 4005–4016.

Méndez-Vigo, B., X. F. Picó, M. Ramiro, J. M. Martinez-Zapater, and C. Alonso-Blanco, 2011 Altitudinal and climatic adaptation is mediated by flowering traits and *FRI*, *FLC*, and *PHYC* genes in *Arabidopsis*. Plant Physiol. 157: 1942–1955.

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*. Proc. Natl. Acad. Sci. USA 100: 10102–10107.

Nei, M., 1987 *Molecular Evolutionary Genetic*. Columbia University Press, New York.

Okazaki, K., K. Sakamoto, R. Kikuchi, A. Saito, E. Togashi *et al.*, 2007 Mapping and characterization of *FLC* homologs and QTL analysis of flowering time in *Brassica oleracea*. Theor. Appl. Genet. 114: 595–608.

Ossowski, S., K. Schneeberger, R. M. Clark, C. Lanz, N. Warthmann *et al.*, 2008 Sequencing of natural strains of *Arabidopsis thaliana* with short reads. Genome Res. 18: 2024–2033.

Pires, J. C., J. Zhao, M. E. Schranz, E. J. Leon, P. A. Quijada *et al.*, 2004 Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploids (Brassicaceae). Biol. J. Linn. Soc. 82: 675–688.

Protas, M. E., and N. H. Patel, 2008 Evolution of coloration patterns. Annu. Rev. Cell Dev. Biol. 24: 425–446.

Quattrocchio, F., J. Wing, K. van der Woude, E. Souer, N. de Vetten *et al.*, 1999 Molecular analysis of the *anthocyanin2* gene of petunia and its role in the evolution of flower color. Plant Cell 11: 1433–1444.

Razi, H., E. C. Howell, H. J. Newbury, and M. J. Kearsey, 2007 Does sequence polymorphism of *FLC* paralogues underlie flowering time QTL in *Brassica oleracea?* Theor. Appl. Genet. 116: 179–192.

Rozas, J., J. C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas, 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496–2497.

Salomé, P. A., K. Bomblies, R. A. Laitinen, L. Yant, R. Mott *et al.*, 2011 Genetic architecture of flowering-time variation in *Arabidopsis thaliana*. Genetics 188: 421–433.

Sánchez-Bermejo, E., B. Méndez-Vigo, F. X. Picó, J. M. Martinez-Zapater, and C. Alonso-Blanco, 2012 Novel natural alleles at *FLC* and *LVR* loci account for enhanced vernalization responses in *Arabidopsis thaliana*. Plant Cell Environ 35: 1672–1684.

Scarcelli, N., and P. X. Kover, 2009 Standing genetic variation in *FRIGIDA* mediates experimental evolution of flowering time in *Arabidopsis*. Mol. Ecol. 18: 2039–2049.

Schneeberger, K., S. Ossowski, C. Lanz, T. Juul, A. H. Petersen *et al.*, 2009 SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nat. Methods 6: 550–551.

Schranz, M. E., P. Quijada, S. B. Sung, L. Lukens, R. Amasino *et al.*, 2002 Characterization and effects of the replicated flowering time gene *FLC* in *Brassica rapa*. Genetics 162: 1457–1468.

Schwinn, K., J. Venail, Y. Shang, S. Mackay, V. Alm *et al.*, 2006 A small family of *MYB*-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. Plant Cell 18: 831–851.

Sheldon, C. C., J. E. Burn, P. P. Perez, J. Metzger, J. A. Edwards *et al.*, 1999 The *FLF* MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. Plant Cell 11: 445–458.

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

Shindo, C., C. Lister, P. Crevillen, M. Nordborg, and C. Dean, 2006 Variation in the epigenetic silencing of *FLC* contributes to natural variation in *Arabidopsis* vernalization response. Genes Dev. 20: 3079–3083.

Slotte, T., H. R. Huang, K. Holm, A. Ceplitis, K. S. Onge *et al.*, 2009 Splicing variation at a *FLOWERING LOCUS C* homeolog is associated with flowering time variation in the tetraploid *Capsella bursa-pastoris*. Genetics 183: 337–345.

Smith, S. D., and M. D. Rausher, 2011 Gene loss and parallel evolution contribute to species difference in flower color. Mol. Evol. Biol. 28: 2799–2810.

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

Strange, A., P. Li, C. Lister, J. Anderson, N. Warthmann *et al.*, 2011 Major-effect alleles at relatively few loci underlie distinct vernalization and flowering variation in *Arabidopsis* accessions. PLoS ONE 6: e19949.

Streisfeld, M. A., and M. D. Rausher, 2009 Genetic changes contributing to the parallel evolution of red floral pigmentation among *Ipomoea* species. New Phytol. 183: 751–763.

Tadege, M., C. C. Sheldon, C. A. Helliwell, P. Stoutjesdijk, E. S. Dennis *et al.*, 2001 Control of flowering time by *FLC* orthologues in *Brassica napus*. Plant J. 28: 545–553.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins, 1997 The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4876–4882.

Toomajian, C., T. T. Hu, M. J. Aranzana, C. Lister, C. Tang *et al.*, 2006 A nonparametric test reveals selection for rapid flowering in the *Arabidopsis* genome. PLoS Biol. 4: e137.

Uptmoor, R., J. Li, T. Schrag, and H. Stutzel, 2012 Prediction of flowering time in *Brassica oleracea* using a quantitative trait loci-based phenology model. Plant Biol. 14: 179–189.

Wang, N., W. Qian, I. Suppanz, L. Wei, B. Mao *et al.*, 2011 Flowering time variation in oilseed rape (*Brassica napus* L.) is associated with allelic variation in the *FRIGIDA* homologue *BnaA.FRI.a*. J. Exp. Bot. 62: 5641–5658.

Wang, R., S. Farrona, C. Vincent, A. Joecker, H. Schoof *et al.*, 2009 *PEP1* regulates perennial flowering in *Arabis alpina*. Nature 459: 423–427.

Weigel, D., 2012 Natural variation in Arabidopsis: from molecular genetics to ecological genomics. Plant Physiol. 158: 2–22.

Weigel, D., and J. Glazebrook, 2002 *Arabidopsis: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

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

Whittall, J. B., C. Voelckel, D. J. Kliebenstein, and S. A. Hodges, 2006 Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in *Aquilegia*. Mol. Ecol. 15: 4645–4657.

Wood, T. E., J. M. Burke, and L. H. Rieseberg, 2005 Parallel genotypic adaptation: when evolution repeats itself. Genetica 123: 157–170.

Ye, K., M. H. Schulz, Q. Long, R. Apweiler, and Z. Ning, 2009 Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics 25: 2865–2871.

Yuan, Y. X., J. Wu, R. F. Sun, X. W. Zhang, D. H. Xu *et al.*, 2009 A naturally occurring splicing site mutation in the *Brassica rapa FLC1* gene is associated with variation in flowering time. J. Exp. Bot. 60: 1299–1308.

Zhao, J., A. Artemyeva, D. P. Del Carpio, R. K. Basnet, N. Zhang *et al.*, 2010a Design of a *Brassica rapa* core collection for association mapping studies. Genome 53: 884–898.

Zhao, J., V. Kulkarni, N. Liu, D. P. Del Carpio, J. Bucher *et al.*, 2010b *BrFLC2* (*FLOWERING LOCUS C*) as a candidate gene for a vernalization response QTL in *Brassica rapa*. J. Exp. Bot. 61: 1817–1825.

*Communicating editor: J. Borevitz*

GENETICS

**Supporting Information**
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.112.143958/-/DC1

# Independent *FLC* Mutations as Causes of Flowering-Time Variation in *Arabidopsis thaliana* and *Capsella rubella*

**Ya-Long Guo, Marco Todesco, Jörg Hagmann, Sandip Das, and Detlef Weigel**


DOI: 10.1534/genetics.112.143958

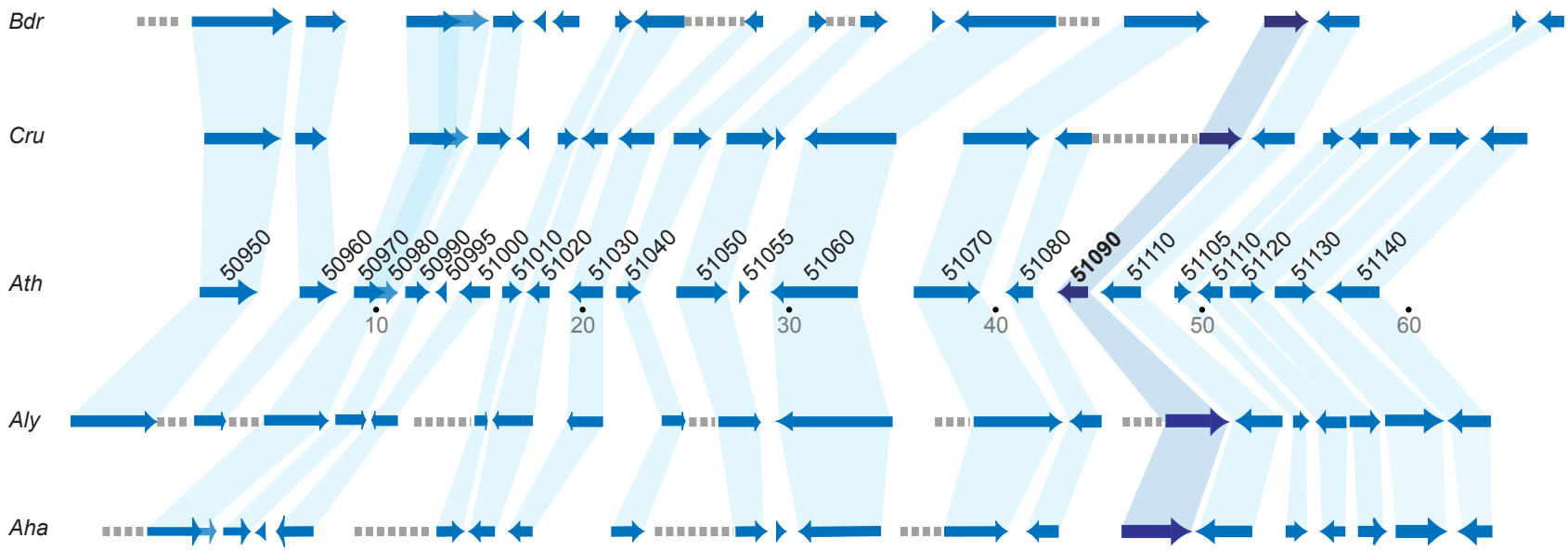


**Figure S1** Comparison of BAC sequences from *Boechera drummondii* (*Bdr*, 93 kb), *Capsella rubella* (*Cru*, 75 kb), *Arabidopsis lyrata*(*Aly*, 110 kb), and *A. halleri* (*Aha*, 92 kb) to a syntenic region in *A. thaliana* (*Ath*, 60 kb), from which *A. thaliana FRIGIDA* (*FRI*) originated. Black numbers indicate At5gXXXXX gene identifiers. The *FRI* homolog, which is truncated in *A. thaliana*, is indicated in bold (see also IRWIN, J. A., C. LISTER, E. SOUMPOUROU, Y. ZHANG, E. C. HOWELL *et al.*, 2012 Functional alleles of the flowering time regulator *FRIGIDA* in the *Brassica oleracea* genome. BMC Plant Biol. **12:** 21). Grey numbers indicate distances in *A. thaliana* in kb. Dashes indicate non-syntenic sequences.

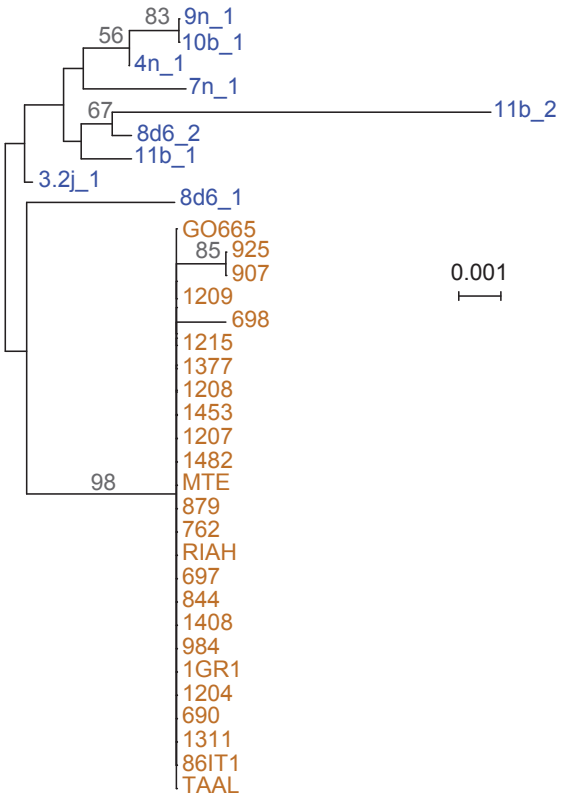


**Figure S2** Phylogenetic tree of *FLC* alleles from *C. rubella* (blue) and *C. grandiflora* (ochre), constructed using PAUP* version 4.0b10 (SWOFFORD, D. L., 2003 PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts) and the Kimura-2 parameter model (KIMURA, M., 1980 A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16:** 111-120; SAITOU, N., and M. NEI, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406-425). Topological robustness was assessed by bootstrapping with 1,000 replicates (FELSENSTEIN, J., 1985 Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:** 783-791). Bootstrap report above 50% is indicated in gray. Scale indicates frequency of substitutions.

**Table S1 Oligonucleotide primers.**

| Purpose | Primer | Strand | Sequence |
|---|---|---|---|
| | | | ***Capsella rubella*** |
| *FLC* gDNA | N-1075 | + | GAA TTC ATG GGA AGA AAA AAA CTA GAA ATC AAG CG |
| | G-9544 | - | ACA TGG TTT TGG ATT TCT GG |
| | G-9759 | - | ACC ATA GTT CAG AGC TTT TGA CTG |
| | G-9854 | + | TCT GAT GCG TGC TCG ATG TTG |
| | G-9898 | + | ATG TTG AAG CTT GTT GAG AA |
| | G-10014 | + | AAA TGT TTC TTC TGC CAT GC |
| | G-10015 | - | TTC AGC AGG TTG AAA ATG ACA |
| | G-10040 | + | TGA CAA TTG ACA ACC CTC CA |
| | G-10228 | - | AGC CGG TCT TCC ATT TTG TA |
| | G-10287 | + | TCT TAA AGC CTT GGT AAT ACA AAC A |
| | G-10320 | - | CCA ATG ATC AAC ACT ACA ATG TCA |
| | G-10443 | + | CCA CTC CTT TTT ATG GAT TTG C |
| | G-10997 | + | ATT CGG TCT GGT TTG AGT TGA G |
| | G-10998 | + | AAG TTT ACG GCT GTG TTT CCA T |
| | G-11096 | - | AAG ATC ACC ATG TTC AGC AAT CA |
| | G-11116 | + | TCT CTG GAA AGA ACC TTG TCC T |
| | G-11218 | + | CCA TGT CAT TAG GTT GGG GTT A |
| | G-11219 | - | AAA TTC AAA CCC GTT CAA TCA T |
| | G-11297 | + | GAT TTA TCG TAG TTT TGT TAT CCA |
| | G-31703 | + | ATG GGG AGA AAA AAA CTA GAA ATC A |
| | G-31704 | - | CTA ATT AAG CAG CGG GAG AGT CAC |
| *FLC* cDNA | G-30600 | + | GAG GAT CAA ATT AGG GCA CAA G |
| | G-30601 | - | CAT GGT TTT GGA TTT CTG GTT T |
| *FLC* splicing analysis | G-12710 | + | GAG ACC GCC CTT TCT GTA ACT A |
| | G-30481 | - | GGA AGA TTG TCG GAG ATT TGT C |
| *FRI* gDNA | G-9537 | - | TGA TGT ATC TGA GGT TGA CTA |
| | G-9748 | + | CTT TCA AAC GCC AAT TCG ATG AT |
| | G-9749 | - | CCC AAA TAT CTT TCT TCA GAT GG |
| | G-9884 | + | CAG ATG GAG AAC TTT TAA TTA GGG |
| | G-9905 | + | TGA AGG AGG ATT AGC TGT GGC |
| | G-9906 | - | TCG TCT CTT TGA CTA GGA AAG |

| | | | |
|---|---|---|---|
| *FLC* qRT-PCR | G-12710 | + | GAG ACC GCC CTT TCT GTA ACT A |
| | G-12711 | + | CAG GTG ACA TCT CCA TCA TCT C |
| *BETA-TUBULIN2* qRT-PCR | G-12712 | - | AGC TTG TTG AGA ATG CTG ATG A |
| | G-12713 | + | GGT CAC CAA AGC TAG GGG TAG T |
| | | | ***Arabidopsis thaliana*** |
| *FLC* gDNA | G-32190 | + | ATG GGA AGA AAA AAA CTA GAA ATC AA |
| | G-32191 | - | CTA ATT AAG TAG TGG GAG AGT CAC |
| *FLC* qRT-PCR | G-30480 | + | TGA GAA CAA AAG TAG CCG ACA A |
| | G-32817 | - | CCG GAG GAG AAG CTG TAG A |
| *FT* qRT-PCR | G-30966 | + | CCC TGC TAC AAC TGG AAC AAC |
| | G-30967 | - | CAC CCT GGT GCA TAC ACT G |
| *SOC1* qRT-PCR | G-30974 | + | ACG AGA AGC TCT CTG AAA AG |
| | G-30975 | - | GAA CAA GGT AAC CCA ATG AAC |
| *BETA-TUBULIN2* qRT-PCR | N-0078 | + | GAG CCT TAC AAC GCT ACT CTG TCT GTC |
| | N-0079 | - | ACA CCA GAC ATA GTA GCA GAA ATC AAG |

**Table S2 SNPs in accession 1408 used as SHOREmap markers.**

Table S2 is available for download as a Gzip compressed archive file at
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.112.143958/-/DC1.

Column 1: sample ID

Column 2: chromosome

Column 3: position

Column 4: reference base

Column 5: alternative base

Column 6: SHORE quality score (between 0 and 40, 40 best)

Column 7: read support of alternative base

Column 8: concordance of alternative base

Column 9: average number of equal-best alignments of alternative base-supporting reads (denoted as repetitiveness)

**Table S3 Flowering times of 305 $F_2$ plants used for SHOREmapping.**

Orange indicates plants that were pooled for bulked segregant analysis.

| Plant ID | DTF |
|---|---|
| 1 | 98 |
| 2 | 97 |
| 3 | 91 |
| 4 | 86 |
| 5 | 50 |
| 6 | 96 |
| 7 | 45 |
| 8 | 106 |
| 9 | 90 |
| 10 | 93 |
| 11 | 91 |
| 12 | 95 |
| 13 | 91 |
| 14 | 96 |
| 15 | 97 |
| 16 | 92 |
| 17 | 74 |
| 18 | 111 |
| 19 | 89 |
| 20 | 87 |
| 21 | 83 |
| 22 | 104 |
| 23 | 100 |
| 24 | 43 |
| 25 | 44 |
| 26 | 40 |
| 27 | 86 |
| 28 | 90 |
| 29 | 51 |
| 30 | 76 |
| 31 | 87 |
| 32 | 89 |
| 33 | 39 |
| 34 | 79 |
| 35 | 90 |
| 36 | 84 |
| Plant | DTF |

| ID | |
|---|---|
| 37 | 46 |
| 38 | 91 |
| 39 | 101 |
| 40 | 45 |
| 41 | 77 |
| 42 | 88 |
| 43 | 97 |
| 44 | 77 |
| 45 | 73 |
| 46 | 75 |
| 47 | 46 |
| 48 | 45 |
| 49 | 44 |
| 50 | 83 |
| 51 | 71 |
| 52 | 82 |
| 53 | 89 |
| 54 | 94 |
| 55 | 81 |
| 56 | 75 |
| 57 | 103 |
| 58 | 76 |
| 59 | 93 |
| 60 | 44 |
| 61 | 91 |
| 62 | 94 |
| 63 | 93 |
| 64 | 45 |
| 65 | 87 |
| 66 | 91 |
| 67 | 76 |
| 68 | 81 |
| 69 | 105 |
| 70 | 86 |
| 71 | 83 |
| 72 | 41 |
| Plant ID | DTF |

| | |
|---|---|
| 73 | 97 |
| 74 | 101 |
| 75 | 100 |
| 76 | 80 |
| 77 | 83 |
| 78 | 94 |
| 79 | 95 |
| 80 | 86 |
| 81 | 93 |
| 82 | 79 |
| 83 | 90 |
| 84 | 89 |
| 85 | 82 |
| 86 | 46 |
| 87 | 77 |
| 88 | 80 |
| 89 | 123 |
| 90 | 87 |
| 91 | 44 |
| 92 | 90 |
| 93 | 91 |
| 94 | 84 |
| 95 | 101 |
| 96 | 76 |
| 97 | 110 |
| 98 | 91 |
| 99 | 88 |
| 100 | 99 |
| 101 | 46 |
| 102 | 83 |
| 103 | 45 |
| 104 | 97 |
| 105 | 100 |
| 106 | 86 |
| 107 | 96 |
| 108 | 45 |
| Plant ID | DTF |
| 109 | 91 |

| | |
|---|---|
| 110 | 83 |
| 111 | 43 |
| 112 | 106 |
| 113 | 97 |
| 114 | 101 |
| 115 | 90 |
| 116 | 97 |
| 117 | 91 |
| 118 | 81 |
| 119 | 83 |
| 120 | 100 |
| 121 | 84 |
| 122 | 97 |
| 123 | 90 |
| 124 | 85 |
| 125 | 92 |
| 126 | 83 |
| 127 | 84 |
| 128 | 84 |
| 129 | 48 |
| 130 | 89 |
| 131 | 95 |
| 132 | 45 |
| 133 | 83 |
| 134 | 85 |
| 135 | 85 |
| 136 | 70 |
| 137 | 44 |
| 138 | 105 |
| 139 | 87 |
| 140 | 40 |
| 141 | 106 |
| 142 | 76 |
| 143 | 80 |
| 144 | 86 |
| Plant ID | DTF |
| 145 | 89 |
| 146 | 77 |

| | |
|---|---|
| 147 | 113 |
| 148 | 115 |
| 149 | 47 |
| 150 | 74 |
| 151 | 79 |
| 152 | 95 |
| 153 | 102 |
| 154 | 98 |
| 155 | 76 |
| 156 | 83 |
| 157 | 78 |
| 158 | 84 |
| 159 | 93 |
| 160 | 82 |
| 161 | 105 |
| 162 | 50 |
| 163 | 49 |
| 164 | 91 |
| 165 | 47 |
| 166 | 90 |
| 167 | 47 |
| 168 | 99 |
| 169 | 120 |
| 170 | 101 |
| 171 | 104 |
| 172 | 118 |
| 173 | 91 |
| 174 | 54 |
| 175 | 101 |
| 176 | 83 |
| 177 | 96 |
| 178 | 84 |
| 179 | 101 |
| 180 | 112 |
| Plant ID | DTF |
| 181 | 47 |
| 182 | 80 |
| 183 | 100 |

| 184 | 101 |
|---|---|
| 185 | 106 |
| 186 | 101 |
| 187 | 90 |
| 188 | 100 |
| 189 | 84 |
| 190 | 83 |
| 191 | 100 |
| 192 | 119 |
| 193 | 82 |
| 194 | 109 |
| 195 | 49 |
| 196 | 84 |
| 197 | 95 |
| 198 | 87 |
| 199 | 97 |
| 200 | 49 |
| 201 | 50 |
| 202 | 98 |
| 203 | 55 |
| 204 | 101 |
| 205 | 130 |
| **Plant ID** | **DTF** |
| 206 | 90 |
| 207 | 91 |

| 208 | 104 |
|---|---|
| 209 | 105 |
| 210 | 104 |
| 211 | 101 |
| 212 | 41 |
| 213 | 52 |
| 214 | 83 |
| 215 | 49 |
| 216 | 82 |
| 217 | 88 |
| 218 | 41 |
| 219 | 43 |
| 220 | 89 |
| 221 | 93 |
| 222 | 90 |
| 223 | 89 |
| 224 | 45 |
| 225 | 82 |
| 226 | 94 |
| 227 | 44 |
| 228 | 59 |
| 229 | 81 |
| 230 | 89 |
| **Plant ID** | **DTF** |
| 231 | 89 |

| 232 | 88 |
|---|---|
| 233 | 95 |
| 234 | 108 |
| 235 | 98 |
| 236 | 91 |
| 237 | 60 |
| 238 | 40 |
| 239 | 84 |
| 240 | 93 |
| 241 | 53 |
| 242 | 105 |
| 243 | 87 |
| 244 | 112 |
| 245 | 45 |
| 246 | 50 |
| 247 | 47 |
| 248 | 88 |
| 249 | 93 |
| 250 | 83 |
| 251 | 91 |
| 252 | 94 |
| 253 | 104 |
| 254 | 97 |
| 255 | 103 |
| **Plant ID** | **DTF** |

| 256 | 108 |
|---|---|
| 257 | 64 |
| 258 | 147 |
| 259 | 38 |
| 260 | 44 |
| 261 | 81 |
| 262 | 82 |
| 263 | 93 |
| 264 | 84 |
| 265 | 131 |
| 266 | 97 |
| 267 | 110 |
| 268 | 124 |
| 269 | 149 |
| 270 | 110 |
| 271 | 55 |
| 272 | 113 |
| 273 | 90 |
| 274 | 105 |
| 275 | 104 |
| 276 | 98 |
| 277 | 96 |
| 278 | 82 |
| 279 | 44 |
| 280 | 97 |
| **Plant** | **DTF** |

| **ID** | |
|---|---|
| 281 | 91 |
| 282 | 82 |
| 283 | 108 |
| 284 | 111 |
| 285 | 41 |
| 286 | 105 |
| 287 | 107 |
| 288 | 134 |
| 289 | 83 |
| 290 | 76 |
| 291 | 51 |
| 292 | 99 |
| 293 | 106 |
| 294 | 42 |
| 295 | 42 |
| 296 | 86 |
| 297 | 91 |
| 298 | 96 |
| 299 | 90 |
| 300 | 48 |
| 301 | 86 |
| 302 | 93 |
| 303 | 84 |
| 304 | 81 |
| 305 | 49 |

**Table S4 Differences between accession 1408 and reference accession MTE in final mapping interval, as determined by SHORE v0.8 and pindel v0.2.4s. The *FLC* gene model is highlighted in bold and red.**

Table S4 is available for download as an MS Excel Open XML file at
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.112.143958/-/DC1.

Sheet 1: SNPs

Sheet 2: short deletions

Sheet 3: short insertions

Sheet 4: SVs

Sheet 5: complex insertions/deletions