Analysis of the Molecular Basis of Flowering Time Variation in Arabidopsis Accessions[1][w]

Silvia Gazzani[2], Anthony R. Gendall[2,3], Clare Lister, and Caroline Dean*

Department of Cell and Developmental Biology, John Innes Centre, Colney Lane, Norwich NR4 7UH, United Kingdom

Allelic variation at the *FRI* (*FRIGIDA*) and *FLC* (*FLOWERING LOCUS C*) loci are major determinants of flowering time in Arabidopsis accessions. Dominant alleles of *FRI* confer a vernalization requirement causing plants to overwinter vegetatively. Many early flowering accessions carry loss-of-function *fri* alleles containing one of two deletions. However, some accessions categorized as early flowering types do not carry these deletion alleles. We have analyzed the molecular basis of earliness in five of these accessions: Cvi, Shakhdara, Wil-2, Kondara, and Kz-9. The Cvi *FRI* allele carries a number of nucleotide differences, one of which causes an in-frame stop codon in the first exon. The other four accessions contain nucleotide differences that only result in amino acid substitutions. Preliminary genetic analysis was consistent with Cvi carrying a nonfunctional *FRI* allele; Wil-2 carrying either a defective *FRI* or a dominant suppressor of *FRI* function; and Shakhdara, Kondara, and Kz-9 carrying a functional *FRI* allele with earliness being caused by allelic variation at other loci including *FLC*. Allelic variation at *FLC* was also investigated in a range of accessions. A novel nonautonomous *Mutator*-like transposon was found in the weak *FLC* allele in Landsberg *erecta*, positioned in the first intron, a region required for normal *FLC* regulation. This transposon was not present in *FLC* alleles of most other accessions including Shakhdara, Kondara, or Kz-9. Thus, variation in Arabidopsis flowering time has arisen through the generation of nonfunctional or weak *FRI* and *FLC* alleles.

The timing of the floral transition has significant consequences for the reproductive success of plants. Plants need to gauge when both environmental and endogenous cues are optimal before undergoing the switch to reproductive development. To achieve this, a complex regulatory network has evolved consisting of multiple pathways that quantitatively regulate a set of genes (the floral pathway integrators) whose activity causes the transition of the meristem to reproductive development (Simpson and Dean, 2002). As plant varieties spread, there is strong selective pressure for changes in flowering time that confer an advantage in that new environment. Arabidopsis accessions show a range of flowering strategies: Some complete their life cycle very rapidly, whereas others adopt a winter annual habit, overwintering vegetatively and flowering in the favorable conditions of spring. The genetic basis of this has been analyzed in a number of studies (Clarke et al., 1995; Jansen et al., 1995; Mitchell-Olds, 1996; Alonso-Blanco et al., 1998).

Allelic variation at *FRI* (*FRIGIDA*) was found to be a major determinant of vernalization requirement and flowering time variation (Napp-Zinn, 1985; Burn et al., 1993; Lee et al., 1993; Clarke and Dean, 1994). *FRI* functions to increase RNA levels of the floral repressor *FLC* (*FLOWERING LOCUS C*), which represses the expression of genes required for the transition to flowering (Michaels and Amasino, 1999; Sheldon et al., 1999). *FLC* RNA levels are reduced during a long period of cold temperature, namely winter, in a process called vernalization. The antagonistic activities of *FRI* and vernalization on *FLC* expression thus result in the winter annual habit. Analysis of the allelic variation at *FRI* showed that the majority of rapid-cycling accessions carry *FRI* alleles with one of two deletions, both predicted to cause loss-of-function (Johanson et al., 2000), indicating that rapid-cycling accessions have evolved independently, at least twice, from winter annual progenitors. Some of these alleles may have evolved more recently than others (Hagenblad and Nordborg, 2002). Extensive variation at *FRI* was also observed in a study of 25 accessions from western Europe, which revealed six novel loss-of function *FRI* alleles and a high degree of polymorphism within exon 1 (Le Corre et al., 2002). These data reinforce the view that *FRI* is under strong selective pressure and is a major target for natural selection of flowering time in Arabidopsis.

We have continued to analyze the molecular variation at *FRI* and have focused on the five early flowering accessions: Cvi (from Cape Verde Island), Shakhdara and Kondara (from Tadjikistan), Kz-9

[1] This work was supported by the Biotechnology and Biological Sciences Research Council (Core Grant to the John Innes Centre and studentship grant to S.G.).

[2] These authors contributed equally to the paper.

[3] Present address: Department of Botany, La Trobe University, Victoria 3086, Australia.

[w] The online version of this article contains Web-only data. The supplemental material is available at http://www.plantphysiol.org.

* Corresponding author; e-mail caroline.dean@bbsrc.ac.uk; fax 44–1603–450025.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.103.021212.

(from Kazakstan), and Wil-2 (from Russia). These have been categorized as early flowering accessions that do not contain either of the *FRI* deletion alleles present in Landsberg *erecta* (L*er*) or Columbia (Col; Johanson et al., 2000). Previously, we had shown that the earliness of Shakhdara was probably due to allelic differences at *FLC* (Johanson et al., 2000). Therefore, we have analyzed allelic variation at *FLC* in a number of accessions including the previously characterized weak allele present in L*er*. We have detected several single-nucleotide polymorphisms plus a 30-bp insertion and a nonautonomous *Mutator*-like transposable element (TE) both within the first intron of *FLC*, a region shown to be important for *FLC* up-regulation (Sheldon et al., 2002). The distribution of these insertions and their correlation with *FLC* activity has been analyzed in a range of accessions.

RESULTS

Sequence Analysis of *FRI* in Early Flowering Accessions

We sequenced approximately 3.6 kb of the genomic region covering *FRI* in five accessions: Cvi, Wil-2, Shakhdara, Kz-9, and Kondara. These had been characterized as early flowering accessions, flowering at <75 d or <10 leaves (Karlsson et al., 1993; Nordborg and Bergelson, 1999) but did not contain either the L*er*- or Col-type deletion alleles (Johanson et al., 2000). The sequenced region contained 573 bp upstream of the ATG putative translation start codon and 870 bp downstream of the putative translation stop codon. The sequences were compared with the same region from the active H51 *FRI* allele (GenBank accession no. AF228499). Within the FRI coding sequence, Cvi (GenBank accession no. AY198404) showed five nucleotide differences in comparison with H51, four in the first and one in the third exon (Fig. 1). Four resulted in amino acid differences, Pro-12 to Thr, Arg-74 to Cys, Asp-167 to Glu, and Lys-377 to Gln, whereas one, in the second half of the first exon, changed amino acid residue Lys-232 to an in-frame translation stop codon (TAA). This would cause premature termination approximately halfway through the protein. The other differences included three nucleotide changes, a three-nucleotide indel 5′ to the coding region and one polymorphism in intron 1. No differences were found in intron 2 or in the 3′-flanking sequence (Fig. 1).

The *FRI* alleles from Shakhdara (GenBank accession no. AY198401), Kz-9 (AY198402), and Kondara (AY198403) were found to encode identical proteins that differed by only one amino acid (amino acid residue Phe-55 to Ile) from the H51 *FRI* allele. This may be a functionally silent substitution because Phe and Ile are both hydrophobic amino acids. Again, in comparison with H51, several polymorphisms were found 5′ to the coding region, one in intron 1 but none in intron 2 or 3′ to the gene.

Two differences were detected between the predicted *FRI* proteins of H51 and Wil-2 (GenBank accession no. AY198405; Fig. 1). These resulted in changes in amino acids Arg-74 to Cys and Asp-167 to Glu in the first exon. The same amino acid differences were also present in the Cvi allele. Depending on whether the Wil-2 *FRI* allele is active or inactive affects the interpretation of the loss-of-function Lys-232 to stop codon mutation in the Cvi allele. If Wil-2 *FRI* is functional, then the in-frame stop codon could be the primary cause of *FRI* loss-of-function in the Cvi accession. Alternatively, if the Wil-2 allele is inactive, the stop codon could have arisen as a sec-

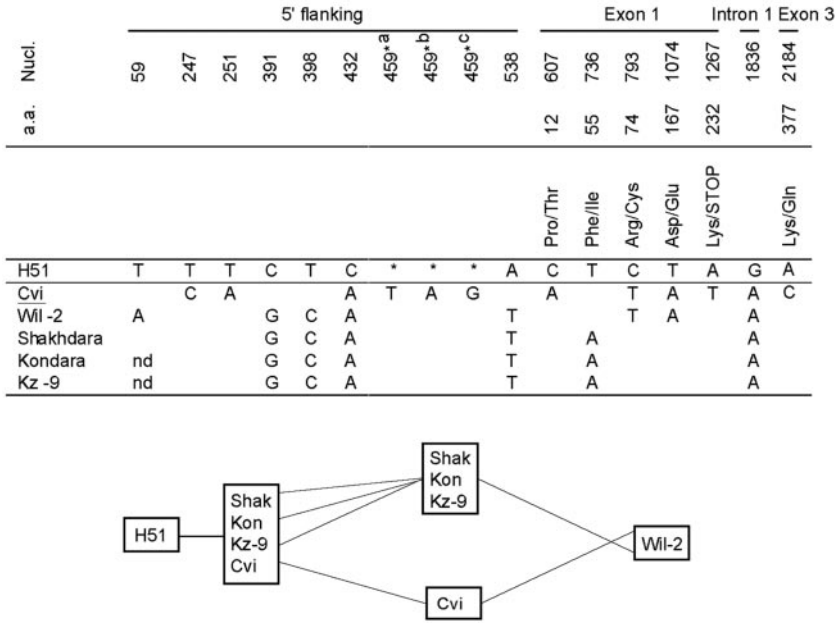

Figure 1. Nucleotide and amino acid changes in the *FRI* gene of the five accessions Cvi, Wil-2, Shakhdara, Kondara, and Kz-9 relative to the active *FRI* allele from H51 (accession no. AF228499). Sequences containing a stop codon are underlined. a.a., Amino acid; Nucl., nucleotide; nd, not determined. The asterisk indicates an insertion. *a, *b, and *c are the first, second, and third nucleotide insertions after nucleotide 459 from H51. These changes have been used to infer the evolutionary relationship between the different *FRI* alleles, and this is represented schematically. The Wil-2 allele is represented as arising from a recombination event between a Cvi-like and a Shakhdara/Kondara/Kz-9-like allele.

ondary mutation due to loss of selection on a nonfunctional coding sequence. The 5′-flanking region of the Wil-2 *FRI* allele showed greater similarity to that of Shakhdara, Kz-9, and Kondara; therefore, the Wil-2 allele may have resulted from a recombination event within the *FRI* gene (Fig. 1).

Genetic Analysis to Test *FRI* Function in the Early Flowering Accessions

To begin to assess whether *FRI* was functional in Cvi, Shakhdara, Kz-9, Kondara, and Wil-2, the flowering time of F_1 seedlings from crosses of the accessions to different backgrounds was determined. This provides a preliminary indication of the activity of the different alleles; however, a definitive conclusion on allele functionality requires more extensive genetic analysis or transformation of the different alleles into a common genetic background. The five accessions were crossed to Col, which carries a nonfunctional *FRI* and a strong *FLC* allele (Koornneef et al., 1994) and in some cases to *FRI (Sf-2)flc-3* (Michaels and Amasino, 1999), which carries an active *FRI* but a nonfunctional *FLC*. The former is most informative for analyzing whether the new accession carries an active *FRI* allele, whereas the latter gives information on the strength of the *FLC* allele. The F_1 plants from the cross between Cvi and Col were early flowering (Table I), supporting a model where Cvi carries an inactive *FRI* allele, consistent with the presence of the in-frame translation stop codon. The F_1 plants from the cross between Shakhdara and Col were late flowering (Table I), most likely due to combining an active *FRI* and a strong *FLC* allele in the F_1 seedlings, thus suggesting that Shakhdara carries a functional *FRI* allele. The major quantitative trait locus (QTL) for lateness mapping near *FRI* in the Shakhdara × Bay-0 recombinant inbred population, therefore, is likely to represent activity from an active Shakhdara *FRI* (Loudet et al., 2002). Preliminary analysis of F_2 populations from a Shakhdara × Col cross had suggested the earliness of Shakhdara was due to a weak (and recessive) *FLC* allele (Johanson et al., 2000), and this is supported by the early flowering of F_1 plants derived from the Shakhdara × *FRI(Sf-2) flc-3* cross (Table I).

Kondara and Kz-9 were not as early flowering in our conditions as had been reported previously (Karlsson et al., 1993; Nordborg and Bergelson, 1999), but the F_1 plants from the crosses to Col flowered even later, suggesting that Kondara and Kz-9 *FRI* alleles are fully functional and confer very late flowering in the presence of the strong Col *FLC* allele. Given the identical amino acid sequence, it is likely that the three accessions Shakhdara, Kondara, and Kz-9 carry the same functional *FRI* allele. The F_1 plants from the cross of Kz-9 to *FRI(Sf2)flc-3* flowered early, supporting the view that, like Shakhdara, Kz-9 earliness is due to a weak *FLC* allele. The Kondara *FRI(Sf2)flc-3* F_1 seedlings were not as early, suggesting that variation at both *FLC* and other loci account for the intermediate flowering time of Kondara. A QTL analysis on F_2 seedlings, F_3 families, or recombinant inbred populations derived from these crosses will be necessary to determine the contribution of the different loci to the flowering time variation.

F_1 plants originating from the cross between Wil-2 and Col flowered early (Table I), suggesting that Wil-2 carries either a nonfunctional *FRI* allele or a dominant suppressor of *FRI* function. Late-flowering individuals (>50 leaves) were not found in the F_2 progeny from this cross (C. Shindo and C. Dean, unpublished data). Because *FRI* can be introduced into Col and confer late flowering (Johanson et al., 2000), Col cannot carry a dominant suppressor of *FRI* function; therefore, these results point to Wil-2 carrying a nonfunctional *FRI* allele or a dominant suppressor being very closely linked to *FRI*. No obvious loss-of-function mutation was present in the Wil-2 allele, and the two amino acid polymorphisms between Wil-2 and H51, Arg-74 to Cys and Asp-167 to Glu, have both been found in *FRI* alleles from late- or very late-flowering accessions, ALL2 and CUR (Le Corre et al., 2002). Further analysis is clearly necessary to determine whether the Arg-74 to Cys and Asp-167 to Glu polymorphisms affect the function of

Table I. *Flowering time analysis*

The flowering time of the indicated genotypes or their F_1 progeny of a cross to Columbia or *FRI(Sf-2)flc-3* (in the Columbia background) was determined by counting total leaf no. Plants were grown in long-day (16 h of light and 8 h of dark) conditions (see "Materials and Methods").

Accession	Flowering Time, Mean ± SE (*n*)		
	Parent	F_1, Col×	F1, *FRI(Sf-2)flc-3*×
Cvi	12.4 ± 0.4 (22)	8.2 ± 0.3 (25)	–
Wil-2	15.4 ± 0.6 (22)	14.0 ± 0.6 (22)	–
Shakhdara	9.3 ± 0.3 (18)	65.2 ± 0.8 (17)	10.4 ± 1.0 (11)
Kz-9	36.6 ± 2.3 (22)	63.0 ± 1.6 (23)	9.4 ± 0.4 (11)
Kondara	57.6 ± 1.3 (21)	70.8 ± 2.7 (20)	18.5 ± 2.5 (27)
L*er*	7.2 ± 0.3 (10)	–	12.8 ± 0.6 (37)
FRI(Sf-2)flc-3	9.3 ± 0.50 (9)	–	–
La0	10.1 ± 0.3 (10)	–	20.2 ± 1.1 (29)
Dijon-G	8.7 ± 0.2 (10)	–	15.3 ± 0.7 (45)

the FRI protein and, if not, why Wil-2 is early flowering.

Analysis of *FRI*/*FLC* RNA Levels

To complement the genetic analysis, we also determined the expression levels of *FRI* and *FLC*. RNA levels and flowering time were compared in the five accessions and the early flowering L*er* and Col genotypes. A fast neutron-induced early flowering *fri* mutant line (FN13; Michaels and Amasino, 2001) and a late-flowering line carrying a cosmid clone containing the active *FRI* allele from the H51 genotype (84M13; Johanson et al., 2000) were also included in the comparison (Fig. 2, A and B). *FRI* is expressed in Cvi, Shakhdara, Kz-9, Kondara, and Wil-2 (Fig. 2B) at a level similar to Col, which is higher than L*er* but lower than lines carrying the H51 *FRI* allele. There does not appear to be a clear correlation between *FRI* RNA levels and flowering time in the five accessions (Fig. 2B). *FLC* RNA levels in general correlate with flowering time, with the notable exception of Cvi.

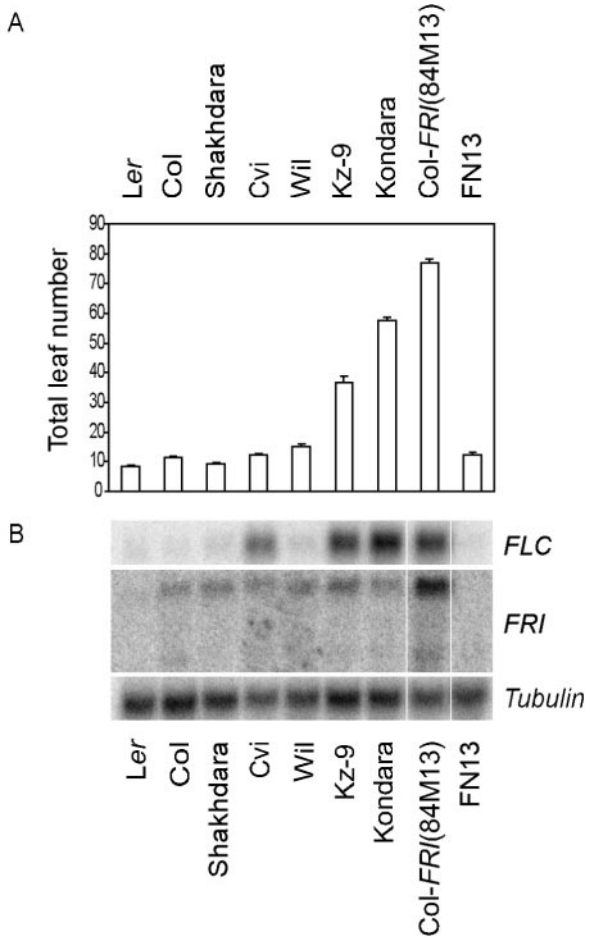

Figure 2. A, Flowering time analysis of the five accessions Cvi, Wil-2, Shakhdara, Kondara, and Kz-9 assayed as total leaf number at flowering. B, Northern analysis of *FRI* and *FLC* in the five accessions Cvi, Wil-2, Shakhdara, Kondara, and Kz-9.

The very early flowering accessions Shakhdara and Wil-2 have low *FLC* RNA levels and the later flowering accessions Kz-9 and Kondara have higher levels, approaching those in lines carrying the active H51 *FRI* allele. The exception is Cvi, which has relatively high *FLC* RNA levels but flowers very early. If Cvi carries a nonfunctional *FRI* allele as suggested by the sequence and genetic analyses, then these results suggest that additional loci other than *FRI* are important in the up-regulation of *FLC* in the Cvi accession. These may include *ART1* (Grbic and Bleecker, 1996), *VIP4* (Zhang and Van Nocker, 2002), *HOS1* (Lee et al., 2001), or the *FLG* QTL, mapped in the Cvi × L*er* recombinant inbred population, which shows a synergistic interaction with a QTL most likely to be *FLC* (Alonso-Blanco et al., 1998).

Analysis of Allelic Variation at *FLC*

Allelic variation at *FLC* does appear to contribute to the flowering time variation in Shakhdara, Kz-9, and Kondara. To date, the most well-characterized *FLC* alleles are from the L*er* and Col accessions (Koornneef et al., 1994; Lee et al., 1994; Sanda and Amasino, 1996). The L*er* *FLC* allele has been described as a weak allele as when combined with active *FRI* alleles lines are significantly earlier than those carrying Col *FLC* alleles (Koornneef et al., 1994; Lee et al., 1994). In the presence of active alleles of *FRI* or loss-of-function mutations of the autonomous pathway *FLC* RNA levels from the L*er* alleles are much lower than *FLC* levels from a Col allele (Schlappi, 2001). C24 and Niederzenz accessions are also considered to carry genetically weak *FLC* alleles (Sanda and Amasino, 1995; Schlappi, 2001). However, genomic blot and sequence analysis of the *FLC* cDNA encoded by the L*er*, Col, C24, and Nd accessions revealed no polymorphism (Sheldon et al., 2000; Schlappi, 2001). We examined *FLC* promoter and intron sequences from the Col (strong *FLC* allele) and C24 (weak *FLC* allele) accessions from position −109 (relative to ATG as +1) and all of intron 1 (accession nos. AL356332 and AF116528, respectively). Single-nucleotide differences were found (positions +217, +515, +1,281, and +2,229), and a 30-bp repeat was present in the first intron of Col but not C24 (Fig. 3, A and B). This repeat does not occur elsewhere in the genome, and it is present in a region required for the maintenance of *FLC* expression after vernalization and for the repression of *FLC* by FCA (Sheldon et al., 2002). We designed a PCR-based marker for this polymorphism and surveyed a total of 23 accessions. This repeat was detected in accessions with strong *FLC* alleles including Col, Li-5, H51 (derived from a Stockholm/Li-5 cross; Napp-Zinn, 1957) and S96 (derived from a Dijon/Li-5 cross; van der Veen, 1965), and also in accessions with weak *FLC* alleles (Niederzenz; Table II; Schlappi, 2001). Therefore, it would not appear to interfere with *FLC* function.

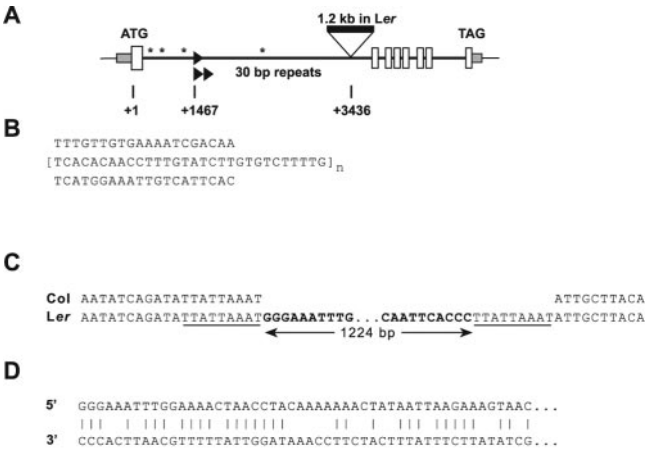

Figure 3. Transposon Insertion in *FLC*. A, Untranslated (gray boxes) and coding regions (white boxes) and start and stop codons (ATG and TAG, respectively) of *FLC* are indicated. The position of the 30-bp repeat is indicated by black arrows, and the insertion site of the transposon in L*er*, DiG, and Di2 is indicated by a triangle. Single-nucleotide substitutions are shown by asterisks. Numbering is based on the A of the ATG start codon as +1. B, Sequence of the 30-bp sequence repeated *n* times, where $n = 1$ or 2. C, Sequence flanking the insertion site in L*er* is shown in plain text, the 9-bp target site duplication (TSD) is underlined, and the transposon sequence is indicated in bold. The Col sequence shown corresponds to bacterial artificial chromosome T31P16 (accession no. AL356332), positions 40,620 to 40,591. D, Inverted repeat of the TE showing the first and last 50 bp of the transposon in L*er*, with complementary residues indicated by vertical lines.

PCR-based analysis of the L*er* allele of *FLC* revealed no major structural changes, apart from one large insertion within intron 1 (Fig. 3, A and C). The sequence of this 1,233-bp insertion was very homologous to three regions of the Col genome cloned in bacterial artificial chromosomes T32N4, MBK23, and F24H14, exhibiting 97%, 94%, and 92% identity over 1,195, 1,242, and 1,187 nucleotides, respectively. It was somewhat less homologous to approximately 17 other regions of the Col genome. It had several characteristics suggesting it was a TE. First, the sequence was delimited by a 9-bp direct repeat, resembling a TSD that occurs upon integration of TEs. The putative TSD sequence 5′-TTATTAAAT is present as a single copy in Col *FLC* (Fig. 3C). Analysis of the homologous sequences revealed that many of them were also flanked by a 9-bp TSD. These TSDs revealed a strong A/T-bias, with only 9% of the nucleotides being G or C, a characteristic feature of *Mutator*-like elements (Yu et al., 2000; for full details, see Supplementary Table I at http://www.plantphysiol.org). Second, the ends of the insertion form an inverted repeat with 32 matches in the first 50 bp (Fig. 3D) and 67 matches over a 120-bp region. Third, the elements share conserved terminal regions: 50% identity over 100 bp at the 5′ end and 22.5% identity over 111 bp at the 3′ end and an additional region of homology (88%–100% identity) between nucleotides 590 and 960.

To assess the prevalence of this putative *Mutator*-like element in different *FLC* alleles, a PCR marker was designed and 27 Arabidopsis accessions analyzed (Table II). The transposon insertion was detected in Dijon-G and Di-2 accessions but not in the closely related Landsberg-0 or Di-1 accessions. The close relationship of Dijon-G and L*er* has been described previously (Koornneef et al., 1994). Analysis of nine additional polymorphic markers failed to identify any differences between L*er* and Di-G (data not shown), implying that they may be the same, or a very closely related, genotype. A number of early flowering accessions, including Shakhdara, do not contain the transposon insertion in the *FLC* locus. If the transposon is the cause of the reduced up-regulation of the L*er FLC* allele, then other mechanisms have led to the weak *FLC* alleles in other Arabidopsis accessions.

DISCUSSION

The molecular analysis of natural variation of Arabidopsis flowering time focused initially on *FRI* because allelic variation at this locus had been shown to be a major genetic determinant of phenotypic variability (Napp-Zinn, 1985; Burn et al., 1993; Lee et al., 1993; Clarke and Dean, 1994). Independent *fri* loss-of-function mutations could account for the earliness of 13 of the 18 rapid-cycling accessions analyzed, suggesting that rapid-cycling accessions had evolved multiple times from winter annual progenitors (Johanson et al., 2000). In this study, we focused on the molecular basis of flowering time variation in the five accessions where the two *fri* deletion alleles did not account for the early flowering behavior. One of these accessions, Cvi, carries an *FRI* allele with an in-frame translation stop codon within exon 1. F_1 analysis suggested the allele was nonfunctional, and this is consistent with the lack of a QTL near *FRI* in the Cvi × L*er* recombinant inbred analysis (Alonso-Blanco et al., 1998). The other accessions show amino acid differences relative to the active H51 *FRI* allele, with three of them, Shakhdara, Kz-9, and Kondara encoding identical, apparently functional, FRI proteins. Further analysis is required to establish whether Wil-2 carries an active *FRI* allele and the functional effects of the Arg-74 to Cys and Asp-167 to Glu substitutions.

The polymorphisms found in the *FRI* alleles described in this study are predominantly in the large first exon of *FRI*. Hypervariability of exon 1 was also found by Le Corre et al. (2002), who argued that the observed excess of non-synonymous polymorphisms and reduced synonymous variation meant that this region of the protein was under strong selection presumably to generate adaptively significant flowering time variation. The first exon contains the first of the two coiled-coil domains, thought to play a role in protein-protein or protein-nucleic acid interactions

Table II. *Molecular variation in FLC*

The presence (+) or absence (–) of the transposon insertion and 30-bp repeat in FLC was detected by PCR- and cleaved-amplified polymorphic sequence-based markers respectively. nd, Not determined. Flowering time categories were as in Karlsson et al. (1993) and Nordborg and Bergelson (1999), but the early flowering category was extended to cover early and intermediate flowering times due to our observations that Kondara and Kz-9 flowered significantly later than Shakhdara, Wil-2, and Cvi.

Flowering Time		Transposon	30-bp Repeat
Early-Intermediate	Late		
Cvi, Shakhdara,Wil-2,WS, LaO, Rsch-4, Gr, Tsu-O, Kz-9, Kondara	H55, St, C24, Öst	–	–
Col, Li5, Nd, S96	H51, Est-O	–	+
Di-1, Te, Can, EDI		–	nd
L*er*, Di-2, Di-G		+	–

(Johanson et al., 2000). However, protein partners of FRI remain to be identified; therefore, interpretation of the consequences of this molecular variability awaits an improved understanding of *FRI* function. The lack of polymorphism in intron 2 and the 3′ region of the gene is striking and may point to purifying selection acting on these regions of the gene. Whether these non-coding sequences have any regulatory function remains to be established.

In addition to continuing with the analysis of molecular variation at *FRI*, we also initiated the analysis of molecular variation at *FLC*. Genetic analysis has suggested that variation at *FLC* also contributes to natural variation in flowering time (Koornneef et al., 1994; Sanda and Amasino, 1996; Johanson et al., 2000), and our results further support this notion. Comparison of the two sequenced *FLC* alleles revealed a 30-bp insertion sequence near the beginning of intron 1, but the presence of this did not appear to be correlated with flowering time. Sequence analysis of the genomic region of *FLC* from L*er*, however, revealed the insertion of a non-TIR *Mutator*-like TE near the 3′ end of intron 1, with the only other changes detected being single-nucleotide changes. The possibility that the transposon is causing the weak *FLC* allele in L*er* is consistent with the observation that the *FLC* cDNAs from L*er* and Col encode identical proteins. Therefore, the weakness of the allele is not due to a polymorphism changing the protein sequence and affecting the function of the protein. How could a transposon positioned within an intron affect a reduced level of RNA in the L*er* genotype compared with Col and a reduced degree of up-regulation by *FRI*? The insertion of transposons within genes can perturb normal gene expression in a number of different ways, particularly by disrupting or reducing the efficiency of normal splicing events or by affecting transcription (for review, see Weil and Wessler, 1990). Thus, the transposon could affect the processing or stability of the primary *FLC* transcript, although there is no evidence of major differences in the pattern of *FLC* transcripts produced by L*er* and other accessions. The transposon insertion in L*er* may lead to reduced transcription of *FLC*, particularly if the insertion interrupts sequences within intron 1 that may normally be involved in the *FRI*-mediated up-regulation. Recent analysis of the cis-elements required for *FLC* regulation has shown that the 3′ region of intron 1(position +3,175 to +3,703), which includes the transposon insertion site in L*er*, is required for both the *FRI*-regulated expression of *FLC* and the repression by vernalization (Sheldon et al., 2002). Alternatively, *Mutator*-like transposons are normally transcriptionally silenced in Arabidopsis through epigenetic modifications (Singer et al., 2001), and these could influence the transcription of the adjacent *FLC*. It will be important to definitively establish that the transposon is the cause of the weak allele through analysis of chimeric *FLC* alleles where the transposon is inserted into a Col *FLC* allele at the same location as it is found in L*er*. We can then pursue the mechanism of how the transposon affects *FLC* regulation.

QTL analysis of Cvi × L*er* recombinant inbred lines had shown that four QTLs (*EDI*, *FLF*, *FLG*, and *FLH*), accounted for the majority of the variation observed (Alonso-Blanco et al., 1998). The *FLF* QTL was considered likely to correspond to *FLC*, and a synergistic interaction of *FLF* and *FLG* suggested that *FLG* functions in a similar way to active *FRI* alleles (Alonso-Blanco et al., 1998). The Cvi *FLG* allele may explain why *FLC* is expressed at relatively high levels in this accession. The high *FLC* levels but early flowering makes Cvi an exception to the general correlation between *FLC* levels and flowering time in Arabidopsis genotypes (Michaels and Amasino, 1999; Sheldon et al., 1999, 2000). The Cvi *EDI* QTL may account for this earliness because it was found to encode a novel allele of *CRYPTOCHROME2 CRY2* (El-Assal et al., 2001). CRY2 is a blue-light photoreceptor that in most accessions promotes flowering in Arabidopsis in long-day photoperiods (El-Assal et al., 2001). In Cvi, the *CRY2* allele confers early flowering in short-day photoperiods, and this enhanced CRY2 activity could function to activate downstream floral pathway integrators, therefore bypassing the repression conferred by high *FLC* levels (Simpson and Dean, 2002).

A number of molecular mechanisms have provided the basis for the evolution of flowering time variation found in Arabidopsis accessions; however, the predominant variation appears to be through changes at

FRI. Further understanding of the molecular control of flowering time and fitness analysis of different genotypes will be required before we can address the question of why so much of the variation in flowering is caused by mutation at *FRI* rather than at other loci. In laboratory mutagenesis experiments, mutations in a large number of loci gave rise to early flowering genotypes. Why has variation at these other loci not been selected for in nature? The timing of the floral transition is probably one of the most important variables in the adaptation of Arabidopsis accessions to new locations. Study of the molecular changes underlying that variation, therefore, is an excellent way to study the molecular basis of adaptation. New Arabidopsis accessions have been or are now being collected with an evolutionary analysis in mind, and these should enable identification of the selective forces driving changes in flowering time.

MATERIALS AND METHODS

Plant Lines

The Arabidopsis accessions (and their Nottingham Arabidopsis Stock Centre nos.) used in this study were Cvi (N1096), Wil-2 (N1596), Shakhdara (N929), Kz-9 (N22441; different batches of this seed have been found to vary in their flowering time), Kondara (N916), L*er* (NW20), Dijon-G (N910), Ws (N1602), H55 (N923), St (N1534), Rsch-4 (N1494), Gr (N1198), Tsu-0 (N1564), C24 (N906), Öst (N1430), Col (N933), Li-5 (N1320), Nd (N1390), S96 (N914), H51 (no stock center no.), Est-0 (N1148), Di-1 (N1108), Te (N1550), Can (N1064), EDI (N1122), and Dijon-2 (N1110).

Crosses of Col to the five accessions Cvi, Wil-2, Shakhdara, Kondara, and Kz-9 were performed. Plants were grown at 20°C with 16 h of light (12 h of fluorescent plus incandescent light at approximately 100 μmol photons m^{-2} s^{-1}, plus 4 h of incandescents only). Crosses of the *FRI(Sf-2)flc-3* line to Shakhdara, Kondara L*er*, La0, and Dijon were carried out in a separate experiment. The F_1 progeny of the crosses and the control plants were grown at 20°C, with a 16-h light photoperiod (fluorescent lights at approximately 100 μmol photons m^{-2} s^{-1}). The flowering time was measured by considering total leaf number, which was scored as the number of rosette leaves plus the number of cauline leaves.

DNA Extraction

Genomic DNA for sequence analysis of the *FRI* allele in the five accessions Cvi, Wil-2, Shakhdara, Kz-9, and Kondara was isolated from young leaves and flowers using a cetyl-trimethyl-ammonium bromide method (Lister et al., 2000). DNA for analysis of the transposon was isolated as described (Dellaporta et al., 1983).

Sequence Analysis of *FRI* Alleles

Overlapping fragments covering the full *FRI* coding sequence, part of the 5′ *FRI*-flanking region (573 bp), and part of the 3′ *FRI*-flanking region (876 bp in Cvi and Shakhdara and 475 bp in Wil-2, Kz-9, and Kondara) were amplified from genomic DNA preparations from the accessions Cvi, Wil-2, Shakhdara, Kz-9, and Kondara. PCR products were purified using the QIAquick Gel Extraction kit (Qiagen USA, Valencia, CA) and used as templates for sequence analysis (for full details, see Supplementary Table II at www.plantphysiol.org). Comparison of at least two independent products of amplification was carried out for each *FRI* fragment. Cycle sequencing reactions were performed using the ABI Prism BigDye Terminator kit (Perkin-Elmer Applied Biosystems, Foster City, CA). Sequences analysis and alignment were performed using the Staden Xgap program (Staden et al., 1998).

RNA Extraction Analysis

Total RNA was extracted from 12 seedlings from each of the five accessions Cvi, Wil-2, Shakhdara, Kondara, and Kz-9 after being grown for 14 d in growth room at 20°C with a 16-h-light/8-h-dark cycle under fluorescent lights at approximately 100 μmol photons m^{-2} s^{-1}. RNA extraction was performed using the TRIzol Reagent (Life Technologies/Gibco-BRL, Cleveland) according to manufacturer's instructions. Total RNA (approximately 10 μg) was fractionated on 1.2% (w/v) formaldehyde-agarose gel and blotted onto Hybond N^+ nylon filters. The RNA gel blot was probed with the full ^{32}P-ATP-labeled *FRI* cDNA. After stripping in boiling 0.5% (w/v) SDS, the same blot was rehybridized with an *FLC* probe, a 403-bp PCR-amplified cDNA fragment lacking MADS domain sequences (corresponding to nucleotides 298–700 of the *FLC* cDNA; Sheldon et al., 1999). Loading was normalized by stripping in boiling 0.5% (w/v) SDS and hybridization with a *β-TUBULIN* coding region probe (Snustad et al., 1992). Blots were exposed to PhosphorImager screens (Molecular Dynamics, Sunnyvale, CA).

Genotyping *FLC* Alleles

Crosses of *FRI(Sf-2)flc-3* to the various accessions were analyzed by isolating DNA from individual F_1 plants and PCR amplification with primers 5′GCAATAGTTCAATCCGTATCG and 5′TTGTAGCTCGTGCGTAGTCTAGCATAG, which amplifies a 917-bp fragment in *FLC* and 813 bp in *flc-3* because *flc-3* has a 104-bp deletion (Michaels and Amasino, 1999). In all cases, two bands of the anticipated sizes were detected, indicating a true cross (corresponding to a heterozygote with a genotype of *FLC flc-3*). To detect the transposon in *FLC*, DNA was PCR amplified with primers 5′AAACAATCTGGACAGTAGAGGCTTAT and 5′CAGGCTGGAGAGATGACAAAA, which amplifies a 527-bp fragment in the absence of the insertion and a 1,741-bp fragment if the insertion is present. The 30-bp intron 1 repeat was detected by PCR amplification using primers 5′AAATGTAAGCCACATTAATTGGGAAA and 5′ATTAAATCATAATTAAGACCAGGAG, and digestion with *Taq*I restriction enzyme, which produces bands of 342, 287, and 42 bp in accessions with a single repeat, and 372, 287, and 42 bp in accessions with two repeats.

Sequencing and Analysis of Mutator-Like Transposon

The 1,741-kb region containing the transposon was amplified from L*er* and Dijon-G using the primers above. PCR products from multiple independent reactions were sequenced directly using primers used for amplification and 5′CATTGGATAACTAATCTTTGAGC, 5′TGTGAGCTTTATTTAGGTTCTAAG, 5′CACTAATCAAATCACTTAACATCC, and 5′TTTTTGTTTATCTAACATGTTTG. Sequences were complied using the SeqMan program within the LASERGENE software (DNASTAR, Inc., Madison, WI) and analyzed by BLAST searches of GenBank at the National Center for Biotechnology Information (Altschul et al., 1997). Regions showing high similarity were further analyzed for the presence of the conserved ends and the presence of TSDs.

ACKNOWLEDGMENTS

We thank Rick Amasino for supplying the *FRI(Sf2)flc-3* line, Thomas Bureau for discussions on transposons, and Mervyn Smith for looking after the Arabidopsis plants.

Received February 4, 2003; returned for revision February 24, 2003; accepted March 28, 2003.

LITERATURE CITED

Alonso-Blanco C, El-Assal SE-D, Coupland G, Koornneef M (1998) Analysis of natural allelic variation at flowering time loci in the Landsberg *erecta* and Cape Verde Islands ecotypes of *Arabidopsis thaliana*. Genetics **149:** 749–764

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389–3402

Burn JE, Smyth DR, Peacock WJ, Dennis ES (1993) Genes conferring late flowering in *Arabidopsis thaliana*. Genetica **90:** 147–155

Clarke JH, Dean C (1994) Mapping *FRI*, a locus controlling flowering time and vernalization response in *Arabidopsis thaliana*. Mol Gen Genet **242:** 81–89

Clarke JH, Mithen R, Brown JK, Dean C (1995) QTL analysis of flowering time in *Arabidopsis thaliana*. Mol Gen Genet **248:** 278–286

Dellaporta S, Wood J, Hicks J (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep **1:** 19–21

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

Grbic V, Bleecker AB (1996) An altered body plan is conferred on *Arabidopsis* plants carrying dominant alleles of two genes. Development **122:** 2395–2403

Hagenblad J, Nordborg M (2002) Sequence variation and haplotype structure surrounding the flowering time locus *FRI* in *Arabidopsis thaliana*. Genetics **161:** 289–298

Jansen RC, van Ooijen JW, Stam P, Lister C, Dean C (1995) Genotype-by-environment interaction in genetic mapping of multiple quantitative trait loci. Theor Appl Genet **91:** 33–37

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

Karlsson BH, Sills GR, Nienhuis J (1993) Effects of photoperiod and vernalization on the number of leaves at flowering in 32 *Arabidopsis thaliana* (Brassicaceae) ecotypes. Am J Bot **80:** 646–648

Koornneef M, Blankestijn-de Vries H, Hanhart C, Soppe W, Peeters T (1994) The phenotype of the some late-flowering mutants is enhanced by a locus on chromosome 5 that is not effective in the Landsberg *erecta* wild-type. Plant J **6:** 911–919

Le Corre V, Roux F, Reboud X (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 H, Xiong L, Gong Z, Ishitani M, Stevenson B, Zhu JK (2001) The *Arabidopsis HOS1* gene negatively regulates cold signal transduction and encodes a RING finger protein that displays cold-regulated nucleocytoplasmic partitioning. Genes Dev **15:** 912–924

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

Lee I, Michaels SD, Masshardt AS, Amasino RM (1994) The late-flowering phenotype of *FRIGIDA* and mutations in *LUMINIDEPENDENS* is suppressed in the Landsberg *erecta* strain of *Arabidopsis*. Plant J **6:** 903–909

Lister C, Anderson M, Dean C (2000) Genetic mapping using recombinant inbred lines. *In* Z Wilson, ed, *Arabidopsis*: a Practical Approach. Oxford University Press, Oxford, pp 51–75

Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F (2002) Bay-0 × Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in *Arabidopsis*. Theor Appl Genet **104:** 1173–1184

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

Michaels SD, Amasino RM (2001) Loss of *FLOWERING LOCUS C* activity eliminates the late-flowering phenotype of *FRIGIDA* and autonomous pathway mutations but not responsiveness to vernalization. Plant Cell **13:** 935–941

Mitchell-Olds T (1996) Genetic constraints on life-history evolution: quantitative-trait loci influencing growth and flowering in *Arabidopsis thaliana*. Evolution **50:** 140–145

Napp-Zinn K (1957) Untersuchungen zur genetik des kältebedürfnisses bei *Arabidopsis thaliana* L. Heynh. Z Indukt Abstamm Vererbungsl **88:** 253–285

Napp-Zinn K (1985) *Arabidopsis thaliana*. *In* AH Halevy, ed, Handbook of Flowering, Vol 1. CRC Press, Boca Raton, FL, pp 492–503

Nordborg M, Bergelson J (1999) The effect of seed and rosette cold treatment on germination and flowering time in some *Arabidopsis thaliana* (Brassicaceae) ecotypes. Am J Bot **86:** 470–475

Sanda SL, Amasino RM (1995) Genetic and physiological analysis of flowering time in the C24 line of *Arabidopsis thaliana*. Weeds World **2:** 2–8

Sanda SL, Amasino RM (1996) Ecotype-specific expression of a flowering mutant phenotype in *Arabidopsis thaliana*. Plant Physiol **111:** 641–644

Schlappi M (2001) RNA levels and activity of *FLOWERING LOCUS C* are modified in mixed genetic backgrounds of *Arabidopsis thaliana*. Int J Plant Sci **162:** 527–537

Sheldon CC, Burn JE, Perez PP, Metzger J, Edwards JA, Peacock WJ, Dennis ES (1999) The *FLF* MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. Plant Cell **11:** 445–458

Sheldon CC, Conn AB, Dennis ES, Peacock WJ (2002) Different regulatory regions are required for the vernalization-induced repression of *FLOWERING LOCUS C* and for the epigenetic maintenance of repression. Plant Cell **14:** 2527–2537

Sheldon CC, Rouse DT, Finnegan EJ, Peacock WJ, Dennis ES (2000) The molecular basis of vernalization: the central role of *FLOWERING LOCUS C* (*FLC*). Proc Natl Acad Sci USA **97:** 3753–3758

Simpson GG, Dean C (2002) *Arabidopsis*, the Rosetta stone of flowering time? Science **296:** 285–289

Singer T, Yordan C, Martienssen RA (2001) Robertson's *Mutator* transposons in *A. thaliana* are regulated by the chromatin-remodeling gene *Decrease in DNA methylation* (*DDM1*). Genes Dev **15:** 591–602

Snustad DP, Haas NA, Kopczak SD, Silflow CD (1992) The small genome of *Arabidopsis* contains at least nine expressed β-tubulin genes. Plant Cell **4:** 549–556

Staden R, Beal KF, Bonfield JK (1998) The Staden Package: computer methods in molecular biology. *In* S Misener, SA Krawetz, eds, Bioinformatics Methods and Protocols, Vol 132. The Humana Press Inc., Totowa, NJ, pp 115–130

van der Veen JH (1965) Genes for late-flowering in *Arabidopsis*. *In* G Röbbelen, ed, *Arabidopsis* Research, Proceedings of the Gottingen Symposium. Gerd Wasmund and Co., Gelsenkirchen, Germany pp 162–171

Weil CF, Wessler SR (1990) The effects of plant transposable element insertion on transcription initiation and RNA processing. Annu Rev Plant Physiol Plant Mol Biol **41:** 527–552

Yu Z, Wright SI, Bureau TE (2000) *Mutator*-like elements in *Arabidopsis thaliana*: structure, diversity and evolution. Genetics **156:** 2019–2031

Zhang H, Van Nocker S (2002) The *VERNALIZATION INDEPENDENCE 4* gene encodes a novel regulator of *FLOWERING LOCUS C*. Plant J **31:** 663–673