Cis-regulatory changes at *FLOWERING LOCUS T* mediate natural variation in flowering responses of *Arabidopsis thaliana*

Christopher Schwartz*,**†, Sureshkumar Balasubramanian†,‡‡,1, Norman Warthmann*,†, Todd P. Michael*,††, Janne Lempe†,2, Sridevi Sureshkumar†,3, Yasushi Kobayashi†, Julin N. Maloof*,§, Justin O. Borevitz*,‡, Joanne Chory*,§§, and Detlef Weigel*,†,4

*Plant Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037; †Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany; ‡Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637; §Section of Plant Biology, University of California, Davis, California 95616; **Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706; ††Plant Biology and Pathology, Waksman Institute, Rutgers University, Piscataway, New Jersey 08855; ‡‡The University of Queensland, School of Biological Sciences, St. Lucia, Queensland 4072, Australia; §§Howard Hughes Medical Institute, The Salk Institute for Biological Sciences, La Jolla, California 92037

1These authors contributed equally to this work

2Present address: Department of Genome Sciences, University of Washington, Seattle, Washington 98195

3Present address: The University of Queensland, School of Biological Sciences, St. Lucia, Queensland 4072, Australia
Corresponding author: MPI for Developmental Biology, Spemannstraße 37-39, D-72076 Tübingen, Germany. E-mail weigel@weigelworld.org
Running title: Natural variation at FLOWERING LOCUS T

Key words: QTL mapping, flowering time, Arabidopsis thaliana, natural variation, FLOWERING LOCUS T

Contact details of corresponding author:

Detlef Weigel
Department of Molecular Biology
MPI for Developmental Biology
Spemannstraße 37-39
D-72076 Tübingen
Germany
Tel: +49-7071-6011411
Fax: +49-7071-6011412
Email: weigel@weigelworld.org
ABSTRACT

Flowering time, a critical adaptive trait, is modulated by several environmental cues. These external signals converge on a small set of genes that in turn mediate the flowering response. Mutant analysis and subsequent molecular studies have revealed that one of these integrator genes, *FLOWERING LOCUS T (FT)*, responds to photoperiod and temperature cues, two environmental parameters that greatly influence flowering time. As the central player in the transition to flowering, the protein coding sequence of *FT* and its function are highly conserved across species. Using QTL mapping with a new advanced intercross-recombinant inbred line (AI-RIL) population, we show that a QTL tightly linked to *FT* contributes to natural variation in the flowering response to the combined effects of photoperiod and ambient temperature. Using Heterogeneous Inbred Families (HIF) and introgression lines, we fine map the QTL to a 6.7 kb fragment in the *FT* promoter. We confirm by quantitative complementation that *FT* has differential activity in the two parental strains. Further support for *FT* underlying the QTL comes from a new approach, quantitative knockdown with artificial microRNAs (amiRNAs). Consistent with the causal sequence polymorphism being in the promoter, we find that the QTL affects *FT* expression. Taken together, these results indicate that allelic variation at pathway integrator genes such as *FT* can underlie phenotypic variability and that this may be achieved through cis-regulatory changes.
INTRODUCTION

Molecular analysis of the phenotypic variation in life history traits is key to understanding how plants evolve in diverse natural environments. Among such traits, flowering time is critical for the reproductive success of the plant and is highly variable among natural *A. thaliana* strains, providing an attractive paradigm for studying adaptive evolution (HAGENBLAD and NORDBORG 2002; JOHANSON et al. 2000; LEMPE et al. 2005; SHINDO et al. 2005; STINCHCOMBE et al. 2004; WERNER et al. 2005a). Two major environmental parameters that modulate flowering time are light and temperature (KOORNNEEF et al. 1998). Temperature and light conditions vary substantially within the geographical range of *A. thaliana*, and natural populations presumably need to adapt to the local environment to ensure reproductive success. Flowering in *A. thaliana* is generally accelerated by long photoperiods, vernalization (exposure to winter-like conditions) and elevated ambient temperatures (BÄURLE and DEAN 2006). All these cues favor flowering of *A. thaliana* during spring or early summer, although the contribution from each individual cue and the interactions among them vary depending on the local environmental conditions (WILCZEK et al. 2009).

Flowering time is controlled through several genetic cascades that converge on a set of integrator genes including *FLOWERING LOCUS T* (*FT*), which encodes a protein that is highly conserved in flowering plants (AHN et al. 2006; KARDAILSKY et al. 1999; KOBAYASHI et al. 1999). FT and its homologs are very likely an integral part of the mobile signal (florigen) that is produced in leaves and travels to the shoot apex to induce flowering (ABE et al. 2005; CORBESIER et al. 2007; JAEGGER and WIGGE 2007; LIFSCHITZ et al. 2006; LIN et al. 2007; MATHIEU et al. 2007; NOTAGUCHI et al. 2008; TAMAKI et al. 2007; WIGGE et al. 2005). In *A. thaliana*, *FT* expression is controlled by photoperiod, vernalization, and ambient growth temperature. Photoperiod in conjunction with the circadian clock promotes daily oscillations in *FT* RNA levels, which are greatly elevated at the end of long days. The central role of *FT* in determining the
timing of flowering appears to be conserved in many species, making FT an attractive target for altering flowering time in cereals and other plants of economic importance (recently reviewed by KOBAYASHI and WEIGEL 2007; TURCK et al. 2008).

Wild strains of A. thaliana show extensive variation in flowering time and much of this is due to variation in the activity of the floral repressor FLOWERING LOCUS C (FLC). While some of this variation maps to FLC itself, much of it is due to differential activity at the epistatically acting FRIGIDA (FRI) locus (JOHANSON et al. 2000; LEMPE et al. 2005; MICHAELS and AMASINO 1999; MICHAELS et al. 2003; SHELDON et al. 1999; SHINDO et al. 2005; SHINDO et al. 2006). Flowering is typically substantially delayed when the FRI/FLC system is active, unless these plants are first vernalized. However, FRI and FLC do not explain all of the flowering time variation seen in wild strains, and functionally divergent alleles of several additional flowering regulators, including CRYPTOCHROME 2 (CRY2), HUA2, FLOWERING LOCUS M (FLM), PHYTOCHROME C (PHYC) and PHYTOCHROME D (PHYD), have been identified in different strains of A. thaliana (ALONSO-BLANCO et al. 1998; AUKERMAN et al. 1997; BALASUBRAMANIAN et al. 2006a; EL-ASSAL et al. 2001; WANG et al. 2007; WERNER et al. 2005b). Finally, there are many genotype by environment interactions that dramatically affect the contribution of a specific locus to the overall phenotype.

The study of natural variation in A. thaliana has been greatly facilitated through the use of Recombinant Inbred Line (RIL) populations (KOORNNEEF et al. 2004). We have recently established two Advanced Intercross (AI)-RIL sets, in which the genetic map is greatly expanded, allowing for high-resolution QTL mapping (BALASUBRAMANIAN et al. 2009). Here we use one of the new AI-RIL populations along with an independent $F_2$ population to identify the molecular basis of a light and temperature sensitive flowering time QTL that mapped to the promoter of the FT gene. We show that FT is likely the causal gene for variation in light and temperature sensitive flowering. Our results, in combination with those from other species,
suggest that cis-regulatory variation rather than structural variation at FT contributes to phenotypic variation in natural populations.

**MATERIALS AND METHODS**

**Plant material and growth conditions:** Two early flowering accessions, Est-1 (Estland [Estonia]; European Arabidopsis Stock Center N6701) and Col-0 (Columbia; WT-2, Lehle Seeds, Tucson, AZ), were used to create the Al-RIL population for QTL analysis. The population consists of 279 individual lines, genotyped at 221 markers (BALASUBRAMANIAN et al. 2009). A second F2 population, derived from Dra-1 (Drahonin, Czechoslovakia; N1119), a strain that behaves similar to photoperiodic mutants (LEMPE et al. 2005), and Ler (Landsberg erecta; N8581), consisted of 190 F2 plants that were genotyped at 77 markers. The markers and the ft mutant alleles have been previously described (BALASUBRAMANIAN et al. 2009; KOORNNEEF et al. 1991; YOO et al. 2005).

Flowering time was determined under 23°C in long days (LD, 16 hours light/8 hours dark) in growth rooms and under LD conditions in a greenhouse in La Jolla, CA, and under 16°C LD in growth rooms in Tübingen, Germany. Short day (SD, 8 hours light/16 hours dark) experiments and flowering time QTL confirmation studies with HIF plants were conducted in a Percival Series 942 growth chamber at 16°C or 22°C in Madison, WI. For quantitative complementation experiments, F1 populations of Est-1, Col-0, Dra-1, Ler, introgression lines, and other arbitrarily chosen strains crossed with various ft mutants were grown in growth rooms, and flowering time was measured as total leaf number at 16°C LD for four to ten plants. Quantitative complementation and knockdown studies with amiR-ft-1 (SCHWAB et al. 2006) were carried out in Tübingen and Madison in 16°C LD or 23°C LD.

**QTL analyses:** Ten plants per RIL were grown in a completely randomized design and
flowering time was measured as days to flowering and as total number of leaves, which were partitioned into juvenile, adult and cauline leaves. QTL analyses (scanone for simple interval mapping and scantwo for two-dimensional scans) were carried out using the r-qtl package in R (http://r-project.org). QTL significance was determined by permutation testing (1000 runs). The scantwo plots are presented as heat maps of additive and epistatic interactions between markers. A color scale for the LOD scores allows comparison to genome-wide averages. Thermosensitivity, a measure of the response of a particular RIL line to a change from 16°C to 23°C compared to the average response of all RILs, was calculated using the slope of the reaction norms as previously described (LEMPE et al. 2005). Thermosensitivity, which is a quantifiable measure of temperature response, was then used as a trait in QTL mapping.

**HIF experiments:** To confirm the F5I14/FT QTL, two heterogeneous inbred families (HIFs, LOUDET et al. 2005; TUINSTRA et al. 1997) segregating only for the QTL region and derived from RIL 110 and RIL 133 were characterized. Initially, 12 seeds from the S8 generation were genotyped to isolate plants homozygous for each parental allele as well as a plant heterozygous at marker F5I14. Seeds were collected from these plants for subsequent experiments, and 200 segregating progeny were analyzed for an association between flowering time and allele status. SD flowering time was measured using only progeny from the two homozygous lines.

**Fine mapping with NILs:** To fine map the F5I14/FT QTL, we generated near isogenic lines (NILs). Est-1 was crossed to Col-0 and seeds were collected from the F₁ and F₂ generations. F₂ plants were genotyped at F5I14 for the QTL region and at one marker from each of the other four chromosomes. A single plant was selected that was heterozygous at F5I14, and Col-0 hetero- or homozygous at the other markers, and backcrossed to Col-0. This was repeated for the BC₁ to BC₃ generations, selecting for heterozygosity at F5I14 and Col homozygosity for all other regions. From the BC₃F₂ generation, plants that belonged to the
earliest and latest quartile (192 plants) were genotyped at marker F5I14. From this experiment, two lines for each F5I14 allele combination (Est-1/Est-1, Est-1/Col-0, and Col-0/Col-0) were chosen as NILs for subsequent experiments. Progeny testing showed that the flowering behavior of the NILs was stable and that the direction and effect of the alleles agreed with the QTL mapping results. One NIL-Est was genotyped at 94 genome-wide loci, and all but 17 of 182 alleles were Col-0.

**Second round of fine mapping:** To reduce the QTL region further, 700 NIL plants heterozygous at F5I14 were genotyped at two markers (24.1 Mb and 24.6 Mb), identifying 28 plants with a recombination event between the two markers. It was necessary to phenotypically classify each plant by progeny testing because of the relatively small effect of the QTL (15% difference between NIL-Col and NIL-Est). Therefore, 12 progeny from each of 28 recombinants were used to classify each recombinant as early, intermediate, or late flowering. Heterozygous parents were readily apparent due to a relatively large standard deviation of the flowering time of their progeny. Further SNP genotyping combined with DNA sequencing reduced the QTL region to the final interval of 6.7 kb.

**Quantitative complementation and quantitative knockdown:** FT activity was assessed by combining specific natural FT alleles with laboratory-induced ft mutant alleles. ft-1, ft-2 and ft-3 are EMS induced alleles in the Ler background (KOORNNEEF *et al.* 1991), and were crossed to Est-1 and Col-0. ft-10 is a T-DNA insertion line in the Col-0 background (Yoo *et al.* 2005), and was crossed to Ler and Dra-1. Similar experiments were conducted using the two homozygous NILs. Line x cross interaction was determined by the following analysis of variance (ANOVA) model: Total leaf number ~ Line + Cross + Line x Cross. An artificial miRNA, amiR-ft-1, which specifically reduces FT expression (Schwab *et al.* 2006), was introduced into Est-1, Col-0, and the homozygous NILs, and flowering time assayed in 12 or more T1 plants for each genotype. The interaction between the genotype and the presence of the amiR-ft-1 transgene
was assessed by the following ANOVA model: Total leaf number ~ Line + Transgene + Line x Transgene.

**FT expression studies:** Plants were analyzed over a time course as previously described (MICHAEL et al. 2008). Briefly, seeds from each genotype were vapor sterilized, plated on half strength MS 0.8% agar plates and stratified for 4 days at 4°C in the dark. Plates were then transferred to continuous white light at 70 μMol m⁻² s⁻¹ with a daily 12 hrs 22°C/12 hrs 12°C temperature regime and grown for 14 days. On the 15th day, plants were harvested into liquid nitrogen every four hours starting at the transition from 12°C to 22°C, for a total of six time points (0, 4, 8, 12, 16, and 20 hrs). Frozen tissue was disrupted in 2 ml Eppendorf tubes containing three ball bearings using a Retsch (Hann, Germany) shaker. RNA was extracted with RNeasy (Qiagen, Valencia, CA) and 5 μg of RNA was used to prepare cDNA (Invitrogen, Carlsbad, CA). cDNA was diluted 1:20, and FT RNA expression was quantified by quantitative real time PCR (qRT-PCR) with SYBRgreen on a MyIQ system (Biorad, Hercules, CA). qRT-PCR protocol and primers have been described (MOCKLER et al. 2004). The FT primers used were: FT_Q254F, 5'-ATCTCCATTGGTTGGTGACTGATA and FT_Q306R, 5'-GCCAAAGGTTGTCCAGTTGTAG.

**Statistical analysis:** Statistical analysis was carried out using JMP (SAS Institute, Raleigh, NC) or R (http://r-project.org). Student's t-tests as implemented in Microsoft Excel were used to determine the significance of the FT expression difference between NILs.

**Sequencing:** Genomic DNA of two recombinants and the parental accessions, Est-1 (GenBank accession #GQ377110 and #xxxxxxx) and Col-0, were sequenced from 24,327,174 to 24,337,983 bp on chromosome 1. Multiple overlapping fragments of 0.6 to 1.0 kb size were amplified by PCR with Advantage cDNA Polymerase Mix (Clontech, Palo Alto, CA), DNA fragments were purified with a spin column (Qiagen). Sequencing reactions were performed in
house with Big Dye terminator (Applied Biosystems, Foster City, CA). Alignments were generated using MegAlign (DNASTar, Madison, WI). To sequence FT from 24 wild strains, four overlapping fragments covering the coding and upstream region were amplified, purified by gel electrophoresis and sequenced using nested primers on both strands. Sequences were aligned with the ABI Prism 2.1 Autoassembler (Applied Biosystems), and the alignment was manually verified and edited in SeAl (http://tree.bio.ed.ac.uk/).

RESULTS

QTL analysis of flowering behavior for the Est-1 x Col-0 AI-RIL population: Flowering time for the Est-1/Col-0 AI-RIL population was measured under inductive long days (LD) at two different temperatures (16°C and 23°C) and in two different light environments (growth room and greenhouse). The distribution of flowering times, measured as days to flowering or total leaf number, was continuous, and showed in all conditions transgression well beyond the parental values (Figure 1A, Supplemental Table 1). QTL analysis identified a large-effect locus on chromosome 1 that influenced flowering time under all conditions, with the Est-1 allele causing a delay in flowering (Figure 1B). Single marker association with closely linked markers as factors and flowering time as the response revealed that the QTL effect was stronger at 16°C, with the QTL explaining as much as 48% of the total variation in days to flowering, while it accounted for about 25% at 23°C.

The 1.5 LOD confidence interval of this QTL at 16°C in LD spanned a 600 kb region that included the well-known floral regulator, FT (At1g65480) (KARDAILSKY et al. 1999; KOBAYASHI et al. 1999). Twenty-seven RILs carried recombinant chromosomes in this interval. An ANOVA using these 27 lines with flowering time as the response and allelic state at the markers as factors, revealed that the F5I14 marker, which is tightly linked to FT (36 kb away),
had the strongest association among the seven surrounding markers (Supplemental Table 2). In order to confirm the QTL, we compared the flowering times of individuals from two Heterogeneous Inbred Families (HIFs, LOUDET et al. 2005; TUINSTRA et al. 1997) that were segregating for the parental alleles only at this marker. Among 200 segregating progeny for each HIF, plants homozygous for the Est-1 allele flowered three to five days later at 23°C LD and 16°C LD than plants with the Col-0 allele (Figure 1C, D). The differences in flowering time for the homozygous plants were highly significant (p<0.0001), while the heterozygous plants had an intermediate phenotype (Figure 1C, D). In addition, the QTL was specific to LD conditions (Figure 1C, D).

Flowering time data from 23°C LD suggested a second QTL on chromosome 1 at 20.8 Mb (linked to marker nga280), north of F5I14, which is located at 24.3 Mb (Figure 1B). This QTL accounted for about 25% of variation in flowering time, with the Est-1 allele delaying flowering compared to the Col-0 allele (Supplemental Table 3). The effect of the nga280 QTL largely disappeared at 16°C LD, while the F5I14/FT QTL effect became more pronounced. A two-dimensional genome scan revealed distinct QTL interactions at 16°C and 23°C. At 23°C, a significant additive interaction was detected between the two QTL on chromosome 1 (Figure 2A), while at 16°C the F5I14/FT QTL interacted additively with a QTL linked to the marker PLS2 on chromosome 2, for which the Est-1 allele promoted early flowering (Figure 2B, Supplemental Table 3). This QTL was near the PHYB locus, which is known to affect flowering (REED et al. 1993). Since there were two interacting QTL that appeared to modulate thermal responses, we calculated the temperature sensitivity in flowering time for each of the RILs (Supplemental Table 1) and mapped QTL for the same (Figure 2C). This analysis revealed that the F5I14/FT QTL and the PLS2/PHYB QTL on chromosome 2, but not the nga280 QTL on chromosome 1, affected temperature sensitivity.

Co-localizing QTL for flowering time in a Dra-1/Ler F2 population: We have
previously identified Dra-1 as a strain that behaves similarly to photoperiodic mutants in the Col-0 and Ler genetic backgrounds (LEMPE et al. 2005). Dra-1 flowers in short days at a similar time as Col-0 and Ler, but flowers later than these strains in long days (although still earlier than in short days). Analysis of an F2 population derived from a cross between Dra-1 and Ler revealed a continuous distribution of flowering times at 23°C LD. However, at 16°C LD, there was a group of late-flowering plants comprising almost 25% of the population (Figure 2D). Linkage analysis indicated a strong association between flowering time and marker F5I14, with the Dra-1 allele conferring later flowering at 16°C LD. QTL analysis at 16°C LD with 192 F2 plants identified several significant QTL across the genome (Figure 2E). The most robust QTL was centered on F5I14/FT, with the confidence interval overlapping that of the Est-1/Col-0 QTL. In addition, a QTL that co-localized with the PLS2/PHYB QTL detected in the Est-1/Col-0 RIL population was observed, with the Dra-1 allele conferring late flowering. Two QTL were found on chromosome 5, one linked to marker CA72 near the floral repressor FLC, and one linked to MBK5 near the MAF2-4 cluster of FLC homologs (Figure 2E). The Dra-1 allele conferred late flowering at the CA72/FLC QTL, and early flowering at the MBK5/MAF QTL (Supplemental Table 3). A two-dimensional genome scan of the Dra-1/Ler population confirmed the epistasis between the F5I14/FT QTL and the PLS2/PHYB QTL on chromosome 2, and revealed additional additive interactions with other QTL (Figure 2F).

**Fine mapping of the F5I14 QTL in Est-1 x Col-0:** To fine map the F5I14/FT QTL, we introgressed the QTL interval from Est-1 into Col-0. We classified the progeny of 392 descendants of a single BC3 plant, which had been heterozygous at marker F5I14, for extreme flowering behavior. We found that the QTL in this backcross population roughly segregated in a Mendelian manner, with all but four of the 96 latest plants being either homozygous for Est-1 at F5I14 (65 plants) or heterozygous (27 plants). To generate Near Isogenic Lines (NILs), we propagated a single plant for each of the three possible F5I14 genotypes (Col-0/Col-0, Est-
1/Est-1, and Col-0/Est-1). The NIL-Est (homozygous for Est-1 allele at F5I14, with the rest of the genome being mostly Col-0; see MATERIALS AND METHODS) was consistently later than Col-0 or the NIL-Col (Figure 3A). In addition, the NIL-Est flowered later than either parent, suggesting that Est-1 contains alleles at other loci that accelerate flowering compared to the Col-0 allele(s) (Figure 3B, C).

For fine mapping, we identified 28 plants with a recombination event within the 600 kb interval surrounding FT, and phenotyped 12 progeny each from these recombinants. Combining the flowering time information of these 28 families with additional genotyping with SNP markers around FT reduced the QTL to an interval of 9 kb (Figure 3D). The entire 9 kb interval was sequenced in the two last recombinants, which further reduced the QTL interval to a 6.7 kb region upstream of the FT coding region (Figure 3D).

**FT as the causal gene for the F5I14/FT QTL:** Since fine mapping identified a non-coding fragment upstream of FT as the QTL, we tested whether FT was the causal gene underlying the QTL in both the Est-1 x Col-0 and Dra-1 x Ler populations. First, we performed quantitative complementation experiments (LONG et al. 1996) using crosses between strains containing ft mutant alleles and Est-1 or Dra-1. We compared the effect of FT alleles in Est-1 with Col-0, and of Dra-1 with Ler, at 16°C LD, the environment where the QTL effect was strongest in both populations. There was a significant line x cross interaction (p < 0.0001), indicating quantitative non-complementation of the ft mutant alleles by the Est-1 and Dra-1 alleles compared to Col-0 and Ler alleles, respectively (Figure 4A, B).

Second, we performed a similar set of experiments with the NILs. These experiments again confirmed a significant line x cross interaction (p<0.01), with the NIL-Est being unable to fully complement the ft mutant at 16°C LD. We included a number of additional strains as controls in this analysis and observed significant interactions only with Est-1 and Dra-1, and the
respective NILs (Figure 4A, B, Supplemental Figure 1A, B).

Third, we adopted a novel approach, quantitative knockdown, where we tested directly whether inactivating the Est-1 allele of \textit{FT} was less effective in delaying flowering than inactivation of the Col-0 allele, using an artificial microRNA against \textit{FT} (amiRNA-ft-1) (Schwab et al. 2006). We transformed the amiR-ft-1 construct into Est-1, Col-0 and the NILs and analyzed the flowering time of at least 20 independent T1 lines in each background. An ANOVA revealed a significant interaction between genetic background and presence of the transgene, with the flowering time of the NIL-Est being the least affected by knocking down \textit{FT} activity (Figure 4C).

Finally, consistent with the flowering behavior of \textit{ft} loss-of-function mutants, the effects of allelic variation largely disappeared in SD, with the NILs flowering at the same time as Col-0 (Figure 4D). Based on these results, we conclude that \textit{FT} underlies the detected F5I14/FT QTL.

\textbf{Allelic variation leads to \textit{FT} expression differences:} Since the QTL interval did not include any coding sequences, we tested whether the phenotype resulted from a difference in \textit{FT} RNA levels. \textit{FT} expression, which is critically correlated with flowering time, is highest at the end of long days (Imaizumi et al. 2003; YanoVsky and Kay 2002). Since the parental strains, Est-1 and Col-0, flowered at similar times, and the effect of the QTL was modest, we reasoned that the differences in \textit{FT} expression conferred by the two alleles might be small. Therefore, we used conditions where such differences are likely to be most obvious. Temperature is known to affect the circadian clock, which in turn is an important factor in the regulation of \textit{FT} expression (Balasubramanian et al. 2006b; Blázquez et al. 2003; Michael et al. 2008). The clock can be entrained by both light and temperature (Michael et al. 2003), and thermocycles (which are normally composed of cool nights and warm days) affect over 50\% of the \textit{A. thaliana} transcriptome (Michael et al. 2008). Thermocycles have a strong inducing effect on \textit{FT}
expression coupled with causing early flowering (Figure 5A; T.P.M., unpublished). Consistent with the QTL affecting *FT* expression, the NIL-Est had lower *FT* RNA levels compared to the NIL-Col (Figure 5A, B). However, there was little difference in the parents, suggesting that other loci compensate for this allelic variation, which is consistent with the flowering behavior of Est-1 being almost identical to Col-0 at both 16°C LD and 23°C LD.

**Sequence variation at the *FT* locus in *A. thaliana*:** To investigate the basis of the *FT* QTL, we analyzed in more detail sequence diversity at *FT*, by sequencing the coding region in 24 strains with variable flowering times in long days, and by sequencing a 4 kb promoter fragment from 12 strains. We found two large overlapping deletions and several other polymorphisms in the upstream region, but few variants in the coding region. We compared the level of polymorphism in our sequence data with published estimates of genome-wide sequence polymorphisms. The genome-wide average is about five nonsynonymous variants per kb coding sequence among 20 divergent strains of *A. thaliana* (Clark et al. 2007; Nordborg et al. 2005). In the 531 bp *FT* coding region, we found four synonymous changes, but no nonsynonymous changes. In contrast, the non-coding sequence was more variable. Within the final 6.7 kb QTL interval, there were many polymorphisms that differentiated the Est-1 and Col-0 alleles, including a 29 bp and 17 bp deletion in Est-1 relative to Col-0, and one insertion of 10 bp (Supplemental Table 4). Partial sequencing of the Dra-1 promoter region revealed a small number of shared polymorphisms that differentiated Est-1 and Dra-1, strains with less active *FT* alleles, from Col-0 and Ler, strains with more active *FT* alleles.

**DISCUSSION**

**Complex genetic interactions modulate flowering time variation in *A. thaliana*:** In the Est-1/Col-0 AI-RIL population, we have detected at least three distinct QTL that interact in an
environment-dependent manner. The $FT$ QTL has the largest effect, explaining about 20 to 40% of the phenotypic variance depending on the environment, but its impact is modulated by additional loci, as might be expected for a gene that integrates multiple environmental signals. For example, the PLS2/$PHYB$ QTL detected at 16°C on chromosome 2 (Figure 1B, 2B, E) displays significant epistatic interaction with the $FT$ QTL. The effect of the Est-1 allele at the PLS2/$PHYB$ QTL is in the opposite direction of the $FT$ allele of Est-1, providing an explanation for why the two parental strains have very similar flowering times. In addition, our analysis at different temperatures revealed interactions between the two closely linked QTL on chromosome 1. While the $FT$ QTL appeared to be the major factor determining variation at lower temperature, the effect of the linked nga280 QTL increased with higher temperature (M. TODESCO, S.B. & D.W., unpublished). A similar picture could also be seen in the Dra-1 x Ler population, in which multiple QTL were mapped and significant interactions between all QTL were detected. While it is conceivable that many of the QTL effects may eventually be mediated through changes in $FT$ expression levels, these results underscore the complexity of the genetic architecture of flowering time variation in $A. thaliana$.

Some of the complexity might reside within $FT$ itself. The genomic interval surrounding $FT$ is unusual (Figure 3D), with long non-coding regions both upstream and downstream of the $FT$ coding sequence. The $FASCIATA1$ ($FAS1$) gene is 7.3 kb upstream of $FT$ and transcribed in the opposite direction of $FT$, while there are no $bona fide$ open reading frames downstream of $FT$ for over 20 kb. Chromatin structure likely plays a prominent role in the regulation of $FT$, since mutations in $LIKE HETEROCHROMATIN PROTEIN1/TERMINAL FLOWER2$ ($LHP1/TFL2$), which is required for epigenetic silencing, cause ectopic $FT$ expression, as do mutations in $EARLY BOLTING IN SHORT DAYS$ ($EBS$), a gene encoding a putative chromatin remodeling factor (KOTAKE et al. 2003; PIÑEIRO et al. 2003; SUNG et al. 2006; TURCK et al. 2007; ZHANG et al. 2007). An analysis of chromatin identified extensive histone modifications in the 3' region of
These are indicative of epigenetic gene regulation, consistent with a large and complex set of 5' and 3' sequences required for proper FT expression. They likely reflect the integrator function of FT, which is the target of many different pathways affecting flowering time (TURCK et al. 2008).

Flowering-time QTL in the FT region: The analysis of the Dra-1 x Ler and the Est-1 x Col-0 populations, and reports of co-localizing QTL in other populations (EL-LITHY et al. 2006; SHINDO et al. 2006; SIMON et al. 2008; WERNER et al. 2005a), suggest that a QTL near FT contributes to natural variation in flowering behavior of several A. thaliana strains. In total, seven independent QTL mapping experiments (including this work), with nine different strains, identified the FT region as a flowering time QTL when Col-0 and Ler are the common parental strains, although the directionality of the QTL varied. Thus, it needs to be determined whether these QTL reflect linked genes affecting flowering in natural strains, or whether they are indeed due to functionally divergent alleles at FT.

Natural variation at a highly connected gene: By integrating multiple environmental signals, FT has a central position in the genetic network that controls flowering time. FT is expressed predominantly in leaves and it is thought that the small FT protein moves to the shoot apex, where it induces flowering by interacting with the FD transcription factor (ABE et al. 2005; CORBESIER et al. 2007; JAEGGER and WIGGE 2007; LI and DUBCOVSKY 2008; LIFSCHITZ et al. 2006; LIN et al. 2007; MATHIEU et al. 2007; NOTAGUCHI et al. 2008; TAMAKI et al. 2007; WIGGE et al. 2005). A major role for FT in flowering time regulation has been shown in many species, including rice, wheat, and poplar (BÖHLENIUS et al. 2006; KOJIMA et al. 2002; YAN et al. 2006). FT is related in sequence to TERMINAL FLOWER1 (TFL1), which has the opposite effect on flowering (BRADLEY et al. 1997; KARDAILSKY et al. 1999; KOBAYASHI et al. 1999; OHSHIMA et al. 1997). The 537 bp coding sequence of TFL1 has been analyzed previously in a sample of 15 different strains. In this collection, one synonymous change and three non-synonymous
changes were reported, with one of the non-synonymous changes surprisingly affecting a residue that appears to be invariant in the entire TFL1/FT gene family across all flowering plants (OLSEN et al. 2002). In contrast, among a similar size sample of 22 strains, we did not detect a single non-synonymous change in the 531 bp FT coding region. These data suggest that FT is highly conserved at the protein level and more constrained than TFL1, consistent with what has been observed at larger phylogenetic distances (AHN et al. 2006).

The role of cis-regulatory versus coding sequence variation: Among genes responsible for natural variation in A. thaliana flowering, variant alleles at six loci, CRY2, HUA2, FLM, FRI, PHYC, PHYD, are affected in protein activity, which in some cases is absent all together (AUKERMAN et al. 1997; BALASUBRAMANIAN et al. 2006a; EL-ASSAL et al. 2001; GAZZANI et al. 2003; JOHANSON et al. 2000; WANG et al. 2007; WERNER et al. 2005b). The exception is the FRI and HUA2 target FLC, where most alleles are affected in expression levels, and only a minority in protein function (GAZZANI et al. 2003; LEMPE et al. 2005; LIU et al. 2004; SHINDO et al. 2005; SHINDO et al. 2006). FT, which encodes a highly conserved protein (AHN et al. 2006), is the most downstream component of flowering time control for which natural variants have been identified in A. thaliana. Functional variation at FT is associated with expression differences, and in this study the final QTL interval included only regulatory sequences. The identification of the FT QTL increases the slowly growing number of examples where regulatory sequences contribute to natural phenotypic variation in A. thaliana (KLIBENSTEIN et al. 2001; KOORNNEEF et al. 2004; LAMBRIX et al. 2001).

Outside of A. thaliana, differences in expression of the FT orthologs Heading date 3a (Hd3a) in rice and VERNALIZATION3 (VRN3) in wheat are responsible for strain-specific flowering differences in these two grasses (KOJIMA et al. 2002; TAKAHASHI et al. 2009; YAN et al. 2006). In contrast, natural variants at the upstream acting loci Hd1, which encodes the ortholog of CONSTANS (CO) in A. thaliana, and Early heading1 (Ehd1), which has no clear A. thaliana
equivalent, are associated with simple loss-of-function mutations in rice (Doi et al. 2004; Takahashi et al. 2009; Yano et al. 2000), as is the case for the upstream acting wheat VRN2 gene (Yan et al. 2004). Finally, further downstream, regulatory variation has also been identified in natural alleles of VRN1, the wheat ortholog of the FT target APETALA1 (AP1) in A. thaliana (Yan et al. 2003). The theme that emerges from these observations is that downstream factors might be more likely to exhibit regulatory sequence variation, while functional diversification at upstream factors might more often involve changes in protein function. These observations might influence the ongoing debate on the importance of regulatory versus coding sequence variation in adaptation and the evolution of development (Hoekstra and Coyne 2007; Stern and Orgogozo 2008; Wray 2007).

ACKNOWLEDGEMENTS

We thank Rick Amasino for his support of this work, the NSF-supported Arabidopsis Biological Resource Centre (ABRC) and the European Arabidopsis Stock Centre (NASC) for seeds, Oliver Bracko for help with growing plants, Marco Todesco for sharing unpublished information, and Richard Clark, Mark Doyle, Johannes Mathieu, Rachel Mooney, and Daniel Ortiz-Barrientos for comments on the manuscript. This work was supported by NIH NRSA fellowship F23-GM65032-1 (C.S), an EMBO Long-Term Fellowship (S.B.), HFSP and JSPS Postdoctoral Fellowships (Y.K.), NIH grant GM62932 (J.C. and D.W.), the Howard Hughes Medical Institute (J.C.), and the Max Planck Society (D.W.).
REFERENCES


BÖHLENIUS, H., T. HUANG, L. CHARBONNEL-CAMPAA, A. M. BRUNNER, S. JANSSON et al., 2006
CO/FT regulatory module controls timing of flowering and seasonal growth cessation in

BRADLEY, D., O. RATCLIFFE, C. VINCENT, R. CARPENTER and E. COEN, 1997 Inflorescence

CLARK, R. M., G. SCHWEIKERT, C. TOOMAJIAN, S. OSSOWSKI, G. ZELLER et al., 2007 Common
sequence polymorphisms shaping genetic diversity in Arabidopsis thaliana. Science
317: 338-342.

CORBESIER, L., C. VINCENT, S. JANG, F. FORNARA, Q. FAN et al., 2007 FT protein movement
contributes to long-distance signaling in floral induction of Arabidopsis. Science 316:
1030-1033.

DOI, K., T. IZAWA, T. FUSE, U. YAMANOUCHI, T. KUBO et al., 2004 Ehd1, a B-type response
regulator in rice, confers short-day promotion of flowering and controls FT-like gene

EL-ASSAL, S. E.-D., C. ALONSO-BLANCO, A. J. PEETERS, V. RAZ and M. KOORNNEEF, 2001 A QTL
440.

EL-LITHY, M. E., L. BENTSINK, C. J. HANHART, G. J. RUYS, D. ROVITO et al., 2006 New
Arabidopsis recombinant inbred line populations genotyped using SNPWave and their
use for mapping flowering-time quantitative trait loci. Genetics 172: 1867-1876.

GAZZANI, S., A. R. GENDALL, C. LISTER and C. DEAN, 2003 Analysis of the molecular basis of

HAGENBLAD, J., and M. NORDBORG, 2002 Sequence variation and haplotype structure

HOEKSTRA, H. E., and J. A. COYNE, 2007 The locus of evolution: evo devo and the genetics of


Li, C., and J. DUBCOVSKY, 2008 Wheat FT protein regulates VRN1 transcription through interactions with FDL2. Plant J.


LONG, A. D., S. L. MULLANEY, T. F. MACKAY and C. H. LANGLEY, 1996 Genetic interactions between naturally occurring alleles at quantitative trait loci and mutant alleles at candidate loci affecting bristle number in Drosophila melanogaster. Genetics 144: 1497-1510.


EARLY BOLTING IN SHORT DAYS is related to chromatin remodeling factors and regulates flowering in Arabidopsis by repressing FT. Plant Cell 15: 1552-1562.


Yan, L., D. Fu, C. Li, A. Blechl, G. Tranquilli et al., 2006 The wheat and barley vernalization gene VRN3 is an orthologue of FT. Proc. Natl. Acad. Sci. USA 103: 19581-19586.

Yan, L., A. Loukoianov, A. Blechl, G. Tranquilli, W. Ramakrishna et al., 2004 The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303: 1640-1644.


FIGURE 1.—QTL analysis of flowering time in Est-1/Col-0 AI-RIL population. (A) Distribution of flowering time (expressed as leaf number) in the AI-RIL population grown in the greenhouse under 23°C LD, including the means for the Est-1 and Col-0 parents. (B) QTL maps of flowering time (measured as days to flower) under three different growth conditions: blue, 16°C LD in growth room; red, 23°C LD in growth chambers; black, 23°C LD in green house. The horizontal line represents the significance threshold for the LOD score. (C, D) Average flowering times of the HIFs based on the genotype of the F5I14 marker. White bars – homozygous for Col-0; grey bars – heterozygous; black bars – homozygous for Est-1. Significant differences (p<0.0001) between the Est-1 and Col-0 genotypes are indicated by asterisks. Error bars represent standard error of mean (i.e.).
FIGURE 2.—Genetic interactions between flowering time QTL. (A, B) Two-dimensional genome scan in Est-1/Col-0 AI-RIL population in 23°C LD (A) and 16°C LD (B), with epistatic interactions on top, and additive interactions on the bottom. Interactions between markers on each chromosome are shown. Color scale indicates LOD scores for epistatic (left) and additive interactions (right). (C) QTL analyses of thermosensitivity in Est-1/Col-0 AI-RIL population, contrasting flowering at 16°C and 23°C, using three different 23°C datasets: black, growth room 1; red, growth room 2; blue, green house. The thermosensitivity QTL co-localizes with the QTL in the F5I14/FT region (Figure 1B). (D) Distribution of flowering time in Dra-1 x Ler F2 population at 23°C and 16°C LD. (E) QTL analysis of flowering time in Dra-1 x Ler F2 population. (D) Two-dimensional genome scan in Dra-1 x Ler F2 population (see panel A, B for legend).
FIGURE 3.—Fine mapping of the chromosome 1 Est-1/Col QTL. (A) Four-week old NIL-Col. (B) NIL-Est of same age and grown in parallel. (C). Distributions of flowering time for Est-1, Col-0 and the Est-NIL. (D) Fine mapping of the QTL. Transcription units are in purple. The FT gene (At1g65480) is highlighted in yellow. The three levels reflect the progressive rounds of fine-mapping, with the final 6.7 kb mapping interval in the FT promoter shown on the bottom. The flanking markers used for mapping are shown. FAS1 (At1g65470) is the gene to the left.
FIGURE 4.—Genetic evidence for *FT* being causal for chromosome 5 QTL. (A, B) Quantitative complementation assays. (A) Flowering time of F₁ plants from crosses of Dra-1 and Ler to Col-0 (“wild type”) and the isogenic *ft*-10 mutant. (B) Flowering time of F₁ plants from crosses of Est-1 and Col-0 to Ler (“wild type”) and the isogenic *ft*-1 mutant. (C) Quantitative knockdown experiment with artificial miRNA against *FT* (amiR-ft-1) introduced into the two NILs and the two parents. (D) Flowering time of the different genotypes under short days.
FIGURE 5.—Allelic variation affects *FT* expression. (A) Comparison of *FT* expression in Col-0 under thermocycles (12 hours 22°C/12 hours 12°C, continuous white light) and light cycles (16 hours light/8 hours dark, constant 23°C). (B) Comparison of *FT* expression levels in NIL-Col and NIL-Est under thermocycles. Two different lines for each NIL are shown. The second NIL-Est line had very low *FT* expression, and its values are barely visible.