Phenotypic evolution through variation in splicing of the noncoding RNA COOLAIR

Peijin Li, Zhen Tao, and Caroline Dean

John Innes Centre, Norwich NR4 7UH, United Kingdom

The extent to which natural polymorphisms in noncoding sequences have functional consequences is still unknown. A large proportion of the natural variation in flowering in Arabidopsis thaliana accessions is due to noncoding cis polymorphisms that define distinct haplotypes of FLOWERING LOCUS C (FLC). Here, we show that a single natural intronic polymorphism in one haplotype affects FLC expression and thus flowering by specifically changing splicing of the FLC antisense transcript COOLAIR. Altered antisense splicing increases FLC expression via a cotranscriptional mechanism involving capping of the FLC nascent transcript. Single noncoding polymorphisms can therefore be a major contributor to phenotypic evolution through modulation of noncoding transcripts.

Supplemental material is available for this article.

Received January 16, 2015; revised version accepted February 27, 2015.

Understanding the molecular basis of natural phenotypic change is a central question in biology (Alonso-Blanco and Mendez-Vigo 2014). We are studying the variation underpinning the adaptive evolution of Arabidopsis thaliana accessions focusing on flowering time, a trait with considerable fitness consequences (Dittmar et al. 2014). A large proportion of the natural variation in flowering time in the floral repressor locus FLOWERING LOCUS C [FLC] (Shindo et al. 2006). Functionally distinct FLC haplotypes exist in the worldwide population, distinguished by polymorphisms in noncoding regions that influence expression level and the rate of epigenetic silencing of the gene (Li et al. 2014).

FLC regulation involves pathways that link different aspects of chromatin regulation with a set of antisense transcripts collectively called COOLAIR (Szwiezewski et al. 2009, Liu et al. 2010, Castaings et al. 2014, Marquardt et al. 2014). Different FLC expression states are linked to alternative splicing and alternative polyadenylation of COOLAIR [Liu et al. 2007, 2010]. Low FLC expression is associated with use of a small intron and polyadenylation at a proximal region, so-called class I COOLAIR variants [Liu et al. 2007; Hornyk et al. 2010; Marquardt et al. 2014; Wang et al. 2014]. High FLC expression is associated with use of a large intron and polyadenylation at a distal region, the class II variants [Fig. 1]. Cold exposure also influences COOLAIR splicing and polyadenylation [Swiezewski et al. 2009]. This sense–antisense transcriptional circuitry raises the possibility that the noncoding cis polymorphism defining the FLC haplotypes might affect FLC function through changed COOLAIR processing. Here, we show that a single natural noncoding polymorphism can significantly change COOLAIR splicing. This SNP is a major contributor to the functional specialization of one of the FLC haplotypes. The study has significant implications for understanding the molecular basis of phenotypic evolution.

Results and Discussion

A. thaliana accessions show variable COOLAIR processing

We surveyed different FLC haplotypes for altered splicing patterns of COOLAIR [Fig. 1A,B] and identified one haplotype characterized by the accessions Var2-6 and Eden-1 with a distinct COOLAIR class II splicing profile but unchanged class I profile [Fig. 1C,D]. The accessions containing this haplotype are predominantly found in northern Sweden [Fig. 2A, Supplemental Table S1] and are generally later flowering than other A. thaliana accessions [Fig. 2B, Supplemental Table S1], and all consistently show higher than average FLC expression [Li et al. 2014]. We introgressed one representative allele from this haplotype [from the Var2-6 accession] into Columbia and then compared expression levels in the presence and absence of FRIGIDA [FRI], the major up-regulator of FLC expression [Lee et al. 1994; Johanson et al. 2000]. The Var2-6 allele was more strongly expressed than the Col allele in a common fri background, with the relative up-regulation by FRI higher for the Col [17.4-fold] than the Var2-6 [5.2-fold] allele [Fig. 2C,D]. This genetic interaction suggests that cis polymorphism in the Var2-6 FLC allele contributes to a higher expression level, potentially through a mechanism similar to FRI function.

SNP259 affects splicing of COOLAIR

DNA sequence comparison between the Var2-6 and Col-0 FLC alleles revealed 43 polymorphisms. One was in the FLC coding region and introduces a synonymous nucleotide change in exon 7, so the Var2-6 FLC allele encodes a protein identical to Col-0 (Supplemental Table S2). The other polymorphisms were in noncoding regions with one, SNP259 Col-0/Var2-6: G/T, located next to the acceptor splice site of the intron of class II COOLAIR [Fig. 1B, Supplemental Table S2; Liu et al. 2007, 2010; Marquardt et al. 2014]. The functional relevance of SNP259 was explored through analysis of an F2 population from a cross between accessions Ka-0 and Ws-0. These accessions have highly related FLC alleles differing only in SNP259 [Ws-0/Ka-0: G/T] and five other polymorphisms within a 22-base-pair (bp) region [Ws-0/Ka-0: C-AAA/TTTT]. The T nucleotide at SNP259 associated with a high proportion of the natural variation in flowering maps to FLC haplotypes (Shindo et al. 2006). Functionally distinct FLC haplotypes exist in the worldwide population, distinguished by polymorphisms in noncoding regions that influence expression level and the rate of epigenetic silencing of the gene (Li et al. 2014).

© 2015 Li et al. This article, published in Genes & Development, is available under a Creative Commons License [Attribution 4.0 International], as described at http://creativecommons.org/licenses/by/4.0/.
with later flowering [Fig. 2E,F] and a COOLAIR splicing profile similar to that in Var2-6 [Fig. 1D]. These analyses implicated SNP259T in the increased expression of FLC, potentially through an influence on COOLAIR splicing. The SNP259T polymorphism was found to significantly reduce use of the splice acceptor site of the class II intron COOLAIR intron [Fig. 3A] and increase use of an acceptor site 54 bp downstream [Fig. 1B]. This change caused additional alternative splicing of COOLAIR and inclusion of an additional 64-bp exon rarely found in COOLAIR transcripts of the Col haplotype [Fig. 1B]. All of the accessions containing the SNP259T polymorphism showed this additional COOLAIR exon [Fig. 3A]. To confirm that the G259T polymorphism adjacent to the COOLAIR splice acceptor site was the cause of these functional differences, we undertook mutagenesis experiments, reciprocally mutating Col-0 FLC to G259T and Var2-6 FLC to T259G [Fig. 3B]. The Col-G259T mutation changed the pattern of COOLAIR splicing so that it contained the additional exon, resembling Var2-6 FLC [Fig. 3C; Supplemental Figs. S1, S2A]. The Var-T259G changed the pattern of COOLAIR splicing so that it resembled Col FLC [Fig. 3D; Supplemental Figs. S1, S2B]. No difference was detected in FLC sense transcript splicing [Fig. 3E,F]. The G259T mutation caused an increase, and the T259G mutation caused a decrease in FLC expression in transgenic plants [Fig. 4A,B]. Taken together, the results indicate that the single G259T polymorphism contributes to increased FLC expression through specific modulation of COOLAIR class II splicing [Fig. 1B].

To assess whether the G259T polymorphism regulated FLC expression directly as well as indirectly via COOLAIR, we generated the G-to-T mutation at nucleotide 259 in an allele of FLC attenuated in the production of COOLAIR. We previously generated a Col FLC terminator exchange transgene [Col-TEX] in which the COOLAIR promoter had been replaced with an RBCS 3’ terminator region. Production of class I and class II COOLAIR transcripts was disrupted in this transgene, although some unspliced nascent antisense transcript was
still produced [Supplemental Fig. S3; Csurba et al. 2014; Marquardt et al. 2014; Wang et al. 2014]. The G259T polymorphism was introduced into Col-TEX to produce Col-TEX-G259T [Fig. 3B]. Plants carrying Col-TEX-G259T expressed FLC at the same level as and flowered at a time similar to those carrying the wild-type Col-TEX transgene [Fig. 4C; Supplemental Fig. S4], indicating that the G259T up-regulation of FLC expression requires aspects of COOLAIR disrupted in Col-TEX. Since the G259T polymorphism increases FLC expression, we also asked whether it affected vernalization response. Transgenic lines carrying the Var2-6 FLC transgene showed higher FLC expression before, during, and after vernalization compared with lines carrying a Var-T259G FLC transgene [Supplemental Fig. S5]. They also flowered considerably later after 4-wk vernalization [Fig. 4D]. The SNP259T polymorphism thus increases FLC expression in all conditions and as such would confer a requirement for longer vernalization in natural conditions.

**COOLAIR splicing affects the degree of capping of the FLC nascent transcript**

Our earlier genetic analysis suggested that cis polymorphism in the Var2-6 FLC allele might up-regulate FLC

---

**Figure 3.** SNP259 specifically regulates COOLAIR alternative splicing. (A) Accessions with nucleotide SNP259T show reduced COOLAIR class II-1 splicing compared with those containing SNP259G. Splicing efficiency was measured with quantitative PCR (qPCR) using the primers F2R2 and normalized to unspliced COOLAIR transcript levels at the same region. Next, the splicing efficiency was normalized to the value in ColFRISF2. The inset shows the position of primer F2R2 for qPCR analysis of COOLAIR variants in ColFRISF2 and Var-6. Multiple accessions containing different FLC haplotypes are shown [Li et al. 2014]. (B) Schematic illustration of the reciprocal SNP259 constructs in Col, Var-2-6 FLC alleles, and Col FLC with RBCS terminator replacement at the COOLAIR promoter region. FLC and COOLAIR are illustrated at the top. RBCS in black bars shows the replacement of the COOLAIR promoter. The Col-0 and Var-2-6 FLC alleles are indicated with white and grey boxes. (C) G259T mutation in Col FLC caused decreased COOLAIR class II-1 splicing in multiple randomly selected transgenic lines. (D) T259G mutation in Var-2-6 FLC caused increased COOLAIR class II-1 splicing in multiple randomly selected transgenic lines. In C and D, each box plot shows the qPCR data of 10 randomly selected independent transgenic plants for each construct. [E] G259T heterogeneity did not influence FLC sense intron 1 splicing. The Student’s t-tests indicated that the FLC sense intron 1 splicing between Col and Col-G259T or Var and Var-T259G is not significant ([NS]) t < 0.05). [F] G259T heterogeneity did not influence FLC sense intron 6 splicing. Values in A, E, and F are means ± SD from three biological repeats. The Student’s t-tests indicated that the FLC sense intron 6 splicing between Col and Col-G259T or between Var and Var-T259G is not significant ([NS]) t < 0.05).

---

**Figure 4.** G259T heterogeneity affects FLC expression in a COOLAIR-dependent mechanism. (A) The G259T mutagenesis in Col FLC increased FLC expression, as assayed using a pool of 48 independent transgenic lines [Student’s t-test, [*∗∗∗* P < 0.001] and 32 randomly selected independent transgenic lines. (B) T259G mutagenesis in Var-2-6 FLC decreased FLC expression, as assayed using a pool of 48 independent transgenic lines [Student’s t-test, [* * P < 0.05] and 29 randomly selected independent transgenic lines. (C) G259T did not change FLC expression in a COOLAIR terminator exchange construct, as assayed using pools from 21 and 32 independent transgenic lines: Col-TEX and Col-TEX-G259T (Student’s t-test, [NS] P > 0.216). (D) Flowering time comparison of Var/FRI flc-2 and Var-T259G/FRI flc-2 of plants given 4 wk of vernalization. Forty-five and 60 independent transgenic lines were scored, with eight individuals of each line included. The KS test indicated that the difference between transgenic lines is significant [P < 0.001]. In A–D, values are means ± SD from three biological repeats. All of the transgenic plants shown in A–D were in a FRI flc-2 background.
expression through a mechanism similar to FRI function (Fig. 2C,D). FRI increases the proportion of FLC transcripts carrying a 5′ cap [Geraldo et al. 2009], a structure that influences the cotranscriptional fate of nascent transcripts and the stability and translation of spliced transcripts. We compared the 5′ capping of FLC transcripts in different genotypes using an RNA adaptor ligation-mediated PCR assay (Supplemental Fig. S6; Geraldo et al. 2009) and found that a higher proportion of the Var2-6 FLC transcripts have a 5′ cap, similar to the functional effect of FRI (Fig. 5A). Mutation of SNP259 C to T in the Col FLC allele recapitulated the Var2-6 molecular phenotype (Fig. 5B); thus, a single polymorphism (from the 43 that distinguish the Var2-6 from Col FLC) is sufficient to change the proportion of FLC transcripts carrying a 5′ cap. Changes in COOLAIR processing therefore appear to influence FLC expression via a mechanism involving the level of capping of the FLC transcript.

Capping/decapping reactions occur on transcripts in both the cytoplasm and the nucleus (Ramaiah et al. 2012). Analysis of chromatin-bound FLC RNA showed that the SNP259 effect on capping occurs on the FLC nascent transcript in the nucleus (Fig. 5C) and is associated with higher levels of chromatin-bound FLC transcript (Fig. 5D). We interpret these data as suggesting that SNP259 impacts FLC expression through a cotranscriptional mechanism linking antisense splicing with sense strand transcription/capping. Chromatin immunoprecipitation (ChIP) experiments showed that RNA polymerase II (Pol II) levels were approximately twofold higher at the Var2-6 FLC allele introgressed into Col as compared with a Col FLC allele, similar to the effect of FRI on Col FLC (Fig. 5E). The transgenic plants containing the reciprocal SNP259 polymorphisms showed opposing H3K27me3 and H3K36me3 profiles, reinforcing the conclusion that SNP259 influences transcriptional dynamics at FLC (Supplemental Fig. S7A,B, Supplemental Table S3). These data are all consistent with the primary effect of SNP259 resulting in a COOLAIR variant that up-regulates transcription of FLC (or represses less well than other COOLAIR class II variants). How a COOLAIR variant [an antisense transcript] influences expression of FLC remains to be established, but COOLAIR class II variants are associated with chromatin at the 5′ end of FLC [Csorba et al. 2014]. The SNP259 polymorphism may positively influence RNA Pol II dynamics or promote efficient transition to elongation and/or stability of the nascent or spliced transcript, which feeds back to promote transcription (Fig. 5F; Haimovich et al. 2013; Hossain et al. 2013; Bentley 2014).

A single noncoding polymorphism, SNP259, is a major contributor to the functional specialization of one of the FLC haplotypes

The FLC haplotype carrying the SNP259T is predominantly found in accessions that occur in the high latitudes of northern Sweden [Li et al. 2014], where plants experience long cold seasons [Fig. 2A, Supplemental Table S1; climate data from north and south Sweden found at http://www.sweden.climatemps.com]. Higher FLC expression and an increased vernalization requirement likely are to be adaptively important in these regions, enabling plants to avoid precocious flowering before winter is over. This significant change in reproductive timing behavior is the result of a single polymorphism in a non-coding region of a locus. This should provide an important paradigm in the many studies associating polymorphism with phenotypic variation. It will also influence the debate over the different types of molecular variation that underpin adaptive evolution [Halligan et al. 2013].

Materials and methods

Plant materials and growth conditions

Transgenic lines were obtained, and plants grown as described previously [Coustham et al. 2012]. Vernalization was performed at 5°C under short
Li et al.

day (8 h of light) conditions, and the vernalized plants were moved back to warm growth conditions afterward. Flowering time was assayed by days to flower. If the plants had not flowered after 120 d, then the experiment was stopped, and the flowering time was scored as 120 d.

FLC expression analysis

FLC expression analysis was performed as described (Coustham et al. 2012). Extensive transgenic line variability was found when comparing independent lines carrying the same FLC transgene. In order to average this variability and quantify FLC expression, we generated a large number of independent transgenic lines for each transgene and pooled seedlings independently three times for the biological replication for RNA extraction and gene expression analyses.

Association analysis between FLC alleles and flowering time in an F2 population

An F2 population was generated by crossing the accession Ws-0 with Ka-0 and selecting the resulting F1. Two-hundred-eighty-eight F2 individuals were stratified for 3 d at 5°C and grown after no vernalization in long days (16 h light/8 h dark), and bolting time was assayed. The individuals of the F2 population were genotyped with specific PCR markers for Ws-0 and Ka-0 FLC alleles (Forward primer, 5′-GTGTGGTGTGTCATCCTACATCCATGT-3′; Reverse primer, 5′-AAACCCATGCCCCACCTCCGAG-3′) and FRI alleles (Forward primer, 5′-TACAAATGTATTTCATGAAG-3′; Reverse primer, 5′-GGTTGGAATTCGCTTTCATTTTTGTAAG-3′). The association between bolting time and Ws-0 and Ka-0 FLC alleles was confirmed in the two subpopulations containing Ws-0 FRI and Ka-0 FRI respectively. The box plot and statistical analysis in this study were performed in Gene Stat edition 9 with default parameters [Ripatti et al. 2009].

Sequence alignment

FLC sequences were aligned and exported with BioEdit package version 7.0.9.0 with default parameters [Hall 1999].

COOLAIR intron splicing analysis

COOLAIR class I was amplified with primers F1 (5′-TTGTTGTTATTTGGTGGTG-3′ and R1 (5′-CCGGTGGCTTCTTACTAC-3′). Class II was amplified with primers F3 (5′-GTATCCGCGCCGCTGACG-3′) and R1. PCR fragments were assayed in agarose gel, cloned into pGEM-T Easy Vector (Promega) and sequenced. For the spliced FLC class I intron splicing analysis, cDNA was synthesized with the RT primer UBC-RP, F2, and unspliced COOLAIR in distal region LP and quantified with the quantitative PCR (qPCR) primers F2 (5′-CCTTCCCGGGCTAAGTA-3′) and R2 (5′-GTCTCAGGATTAAAGAATGAAA-3′). Unspliced COOLAIR class II intron in the distal region was amplified with primers LP (5′-CGCAATTCTCTCTGATAG-3′) and RP (5′-CTTTGGAATACGCAAGGACGAC-3′). SNPs259/T in class II splicing was quantified with primer F2 and R3 (5′-GAAAGACACCCCGGTTGGAG-3′). The intron splicing was calculated by normalizing spliced sequence to unspliced COOLAIR. The PCR position is shown in Figure 1B. The internal control for qRT-PCR is UBC with primers 5′-CTCGGACTCAGGGAATCTCTTC-3′ and 5′-CTGGACAGTGGAAACC-3′. The sequences of COOLAIR splicing variants in Vae-6 were submitted to GenBank (accession no. KP792767-792771).

FLC sense splicing analysis

We analyzed the intron splicing in FLC introns 1 and 6. For intron 1, 2.5 μg of total RNA was treated with Turbo DNase-free kit [Ambion] and reverse transcribed with primer 5′-CAGGCTCATTTCGACACAGCA-3′. Spliced FLC expression was assayed with primers 5′-GAGAGAAGCCATCGGGGAAGAAGG-3′ and 5′-AGGTTATCGCCGGAGGAG-3′ and unspliced FLC expression was assayed with primers 5′-AAAAAGTGAATTTTGGACGAGAGG-3′ and 5′-GGGAGACTCCCGGAGAATTTG-3′. For intron 6, 2.5 μg of total RNA was transcribed with primer 5′-GGGAGACTCCCGGAGAATTTG-3′. Spliced FLC expression was assayed with primers 5′-GTGGTTGTTATTTGGTGGTG-3′ and 5′-CTCTCATCTCAGCTCTC-3′. The intron splicing efficiency was calculated by normalizing the spliced FLC expression level to the unspliced FLC expression level for intron 1 and intron 6, respectively. The internal control for qRT-PCR was UBC.

Capping analysis of FLC RNA

The capping analysis of spliced FLC RNA was performed using the First-Choice RLM-RACE kit [Life Technologies] following the procedure described previously [Geraldo et al. 2009]. After the first round of PCR, the PCR product was treated with 0.5 U of exonuclease I [New England Biolabs] and 0.25 U of alkaline phosphatase [shrimp] [Roche] for 30 min at 37°C and 5 min at 95°C and used for nested PCR. For the 5′ capping analysis of nascent FLC, the chromatin RNA was prepared as follows: the nuclei were extracted with Honda buffer (0.44 M sucrose, 1.25% Ficoll), 2.5% dextran T40, 20 mM Hepes KOH at pH 7.4, 10 mM MgCl2, 0.5% Triton X-100, 5 mM DTT, 1 mM PMSF, 1× EDTA-free Complete protease inhibitor [Roche], 100 μg/mL yeast tRNA, 40 U/mL Superase In RNase inhibitor [Ambion], 5 mM β-mercaptoethanol] and washed with urea/NP40 buffer (20 mM Tris at pH 8.0, 300 mM NaCl, 7.5 mM MgCl2, 0.2 mM EDTA, 1 mM DTT, 1 M Urea, 1% Igepal CA-630, 1× EDTA-free Complete protease inhibitor [Roche], 20 U/mL Superase In RNase inhibitor [Ambion]). Next, the chromatin pellet was extracted with the hot phenol method for RNA preparation [Box et al. 2011]. 5′ capping analysis was performed using Turbo DNase-treated chromatin RNA following the same procedure. The FLC-specific primers for capping analysis were 5′-CCATACAACTC TATAGACCTC GCTTA-3′ and 5′-CATCGGACCGCATCGAGCTCATG TATCAAAC-3′. The nested PCR products of all samples were cloned into pGEM-T Easy Vector (Promega) and sequenced with M13 forward primer to confirm the sequence identity.

ChIP assay

The histone modification and Pol II occupancy assays were performed as described previously [Wang et al. 2014; Yang et al. 2014]. The qPCR primers for the ChIP assay are included in Supplemental Table S2. For histone modification assays, anti-Histone H3 (trimethyl K27) antibody [Millipore, 07-449], anti-Histone H3 antibody [Abcam, ab1791], and anti-Histone H3 (trimethyl K36) antibody [Abcam, ab9050] were used. For the total Pol II occupancy assay, anti-RNA polymerase II CTD repeat YSPTSPS antibody [JWG16] [Abcam, ab817] was used for ChIP assay.

Acknowledgments

We thank all members of the Dean and Howard groups for discussions. This research was supported by the UK Biotechnology and Biological Sciences Research Council [BBSRC] grant BB/K007857/1, Institute Strategic Programme grant BB/J004588/1 and European Research Council Advanced Investigator grants ENVGENE and MEXTIM.

References


Phenotypic evolution through variation in splicing of the noncoding RNA COOLAIR

Peijin Li, Zhen Tao and Caroline Dean

*Genes Dev.* 2015 29: 696-701 originally published online March 24, 2015
Access the most recent version at doi:10.1101/gad.258814.115

<table>
<thead>
<tr>
<th>Supplemental Material</th>
<th><a href="http://genesdev.cshlp.org/content/suppl/2015/03/27/gad.258814.115.DC1.html">http://genesdev.cshlp.org/content/suppl/2015/03/27/gad.258814.115.DC1.html</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>References</td>
<td>This article cites 25 articles, 9 of which can be accessed free at:</td>
</tr>
<tr>
<td></td>
<td><a href="http://genesdev.cshlp.org/content/29/7/696.full.html#ref-list-1">http://genesdev.cshlp.org/content/29/7/696.full.html#ref-list-1</a></td>
</tr>
<tr>
<td>Open Access</td>
<td>Freely available online through the <em>Genes &amp; Development</em> Open Access option.</td>
</tr>
<tr>
<td>Creative Commons License</td>
<td>This article, published in <em>Genes &amp; Development</em>, is available under a Creative Commons License (Attribution 4.0 International), as described at <a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a>.</td>
</tr>
<tr>
<td>Email Alerting Service</td>
<td>Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or <a href="http://genesdev.cshlp.org/subscriptions">click here</a>.</td>
</tr>
</tbody>
</table>

To subscribe to *Genes & Development* go to: [http://genesdev.cshlp.org/subscriptions](http://genesdev.cshlp.org/subscriptions)

© 2015 Li et al.; Published by Cold Spring Harbor Laboratory Press