Flowering newsletter review

Message ends: RNA 3′ processing and flowering time control

Katarzyna Rataj1 and Gordon G. Simpson1,2,*

1 College of Life Sciences, University of Dundee, Dundee DD1 4HN, UK
2 James Hutton Institute, Invergowrie, Dundee DD2 5DA, UK

* To whom correspondence should be addressed. E-mail: g.g.simpson@dundee.ac.uk

Received 5 September 2013; Revised 22 November 2013; Accepted 26 November 2013

Abstract

Plants control the time at which they flower in order to ensure reproductive success. This control is underpinned by precision in gene regulation acting through genetically separable pathways. The genetic dissection of this process in the model plant Arabidopsis thaliana has led to the recurrent identification of plant-specific and highly conserved RNA 3′ end processing factors required to control flowering by specifically controlling transcription of mRNA encoding the floral repressor FLOWERING LOCUS C (FLC). Here, we review the features of these RNA-processing and RNA-associated proteins, and the complex architecture of coding and non-coding RNA transcription at the FLC locus. We discuss alternative concepts that might explain how these RNA-processing events regulate FLC transcription and hence control flowering time.

Background

By controlling the time at which they flower, plants enhance their chances of reproductive success. This control is responsive to environmental cues like temperature and day length, and internal factors such as age, and thus ensures that flowering occurs under favourable conditions (Andrés and Coupland, 2012). Underpinning this control is precision in gene regulation acting through a set of genetically separable pathways that mediate and integrate these responses to establish flowering. Because of this link between environment and reproductive success, variation in this trait is a key part of how plants adapt to their environment.

The characterization of the molecular basis of flowering time control has revealed that a multiplicity of gene regulatory mechanisms underpin this developmental switch. Prominent among these are changes in gene expression programmes mediated by transcription factors. In addition, mutation of factors that modify chromatin results in altered flowering times. Post-transcriptional and post-translational regulatory events such as protein phosphorylation and ubiquitylation feature in signalling cascades to control transcription factor activity. In addition, microRNA (miRNA)-based regulation of flowering in response to age in annual and perennial flowering species involves miR156- and miR172-mediated control of transcription factors (Aukerman and Sakai, 2003; Huijser and Schmid, 2011; Bergonzì et al., 2013). Alternative splicing of pre-mRNA encoding transcription factors can also impact on flowering time. For example, ambient temperature-dependent alternative splicing of pre-mRNA encoding the floral repressor FLOWERING LOCUS M (FLM) alters its ability to interact with another flowering regulator, SHORT VEGETATIVE PHASE (SVP), and hence control flowering (Lee et al., 2013; Posé et al., 2013).

In this review, we consider the recurrent identification of factors that control 3′ end RNA processing and transcription termination as regulators of flowering time. While these include some apparently plant-specific RNA-binding factors.
proteins, somewhat unexpectedly, very highly conserved splicing and cleavage and polyadenylation factors have repeatedly been identified as affecting flowering in very specific ways.

Contrasting features of flowering time mutants reveal different genetically defined pathways

Our contemporary understanding of the molecular basis of flowering time control derives primarily from a series of breakthroughs made with the model plant Arabidopsis thaliana (Simpson and Dean, 2002). Although not the first to study late-flowering mutants, Maarten Koornneef pioneered the use of A. thaliana mutant screens to dissect complex plant traits in the 1980s and published the characteristics of an especially influential set of late-flowering mutants (Koornneef et al., 1991). In this seminal work, Koornneef classified 11 different late-flowering mutants according to their physiological responses to day length and the long period of cold required to accelerate flowering in some plants that is known as vernalization. He made all possible combinations of late-flowering double mutants to reveal genetic epistasis. From this study, he proposed one of the basic conceptual frameworks for the field – that genetically separable pathways promote flowering. Koornneef defined a pathway promoting flowering in response to day length, which he called the photoperiod pathway, and another that promoted flowering independently of day length, which he first called the constitutive pathway and later renamed as the autonomous pathway.

Koornneef shared his mutants, and through the subsequent combined efforts of the field as a whole, the genes disrupted in each of these late-flowering mutants were identified. Many of their molecular interactions have been studied in detail, and our current understanding of flowering time has been reviewed recently (Andrés and Coupland, 2012). Characterization of the photoperiod response pathway has shown that, fundamentally, it comprises a cascade of transcription factors: the circadian clock is crucial to this pathway, and a flowering-specific output from the clock is the circadian control of transcription of the B-box transcription factor CONSTANS (CO). CO directly activates the transcription of FLOWERING LOCUS T (FT) in phloem companion cells. FT protein then travels through the sieve elements of the phloem to the shoot apex where it interacts with the basic leucine-zipper transcription factor FD to form a complex that directly activates the MADS-box transcription factor and floral meristem identity gene APETALA 1. The Koornneef photoperiod pathway mutants co, ft, and fd therefore flower late because they are unable to actively promote flowering.

In contrast, molecular characterization of the autonomous pathway revealed features quite distinct from those of the photoperiod pathway. First, the genes disrupted in the original Koornneef mutants fca, fpa, and fy (Koornneef et al., 1991). FVE is related to yeast MSI (multicopy suppressor of IRA1) and human retinoblastoma-associated proteins RbAp46 and RbAp48, which are conserved components of chromatin-remodelling complexes (Ausín et al., 2004). In contrast, FCA, FPA, and FY encode proteins that affect cleavage and polyadenylation of mRNA (Quesada et al., 2003; Simpson et al., 2003; Hornyik et al., 2010b).

FCA, an apparently plant-specific protein, is comprised of two RNA-recognition motifs and a WW protein-interaction domain (Macknight et al., 1997). A role for FCA in controlling RNA 3′ end formation was revealed by the simultaneous discoveries that it autoregulated its expression by alternative polyadenylation (Quesada et al., 2003) and that a direct protein-interaction partner necessary for its function, FY, was a conserved core component of the cleavage and polyadenylation machinery (Simpson et al., 2003) [the FY homologue...
in yeast is Pfs2p (Ohnacker et al., 2000) and in humans is WDR33 (Shi et al., 2009).

FCA pre-mRNA is alternatively polyadenylated and FCA promotes the selection of poly(A) sites in the third intron of its own pre-mRNA in a manner dependent on a physical interaction with FY (Quesada et al., 2003; Simpson et al., 2003). The resultant truncated RNA is non-functional in flowering (Macknight et al., 2002). FY is comprised of seven N-terminal WD repeats, and it is this part of the protein that is highly conserved in other eukaryotes. Mutations that disrupt the WD repeats are lethal in yeast (Ohnacker et al., 2000) and in A. thaliana (Henderson et al., 2005), reflecting the essential nature of mRNA cleavage and polyadenylation to gene expression. However, FY has a C-terminal extension not found in yeast, that carries PPLPP sequences conserved in diverse plant species and which interact with the WW domain of FCA (Simpson et al., 2003; Henderson et al., 2005). Mutations of either FCA or FY that disrupt this WW-PPLPP interaction result in late flowering (Simpson et al., 2003; Henderson et al., 2005). The interpretation of these protein interactions led to a model that proposes that the plant-specific RNA-binding protein FCA tethers the cleavage and polyadenylation machinery to a ‘weak’ regulatable poly(A) site that influences flowering time by interacting with a conserved polyadenylation factor FY through a plant-specific protein-interaction domain (Simpson et al., 2003). In this way, a constitutive feature of pre-mRNA maturation can be recruited to the specific control of flowering. Remarkably, natural variation in flowering time in the Bla-6 accession of A. thaliana can be explained by mutation of one of these PPLPP sequences in FY (Adams et al., 2009). In other words, naturally occurring mutations in a core component of the cleavage and polyadenylation machinery can underpin adaptation through flowering time control in the wild.

The characterization of FPA further underlined a connection between the autonomous pathway and RNA 3’ end formation. FPA encodes a protein with three RNA-recognition motifs in the N-terminal region and a SPOC (Spen parologue and orthologue C-terminal) protein-interaction domain at the C-terminus (Schomburg et al., 2001; Hornyik et al., 2010b). This combination of domains defines the signature features of SPEN proteins, a protein family found in a range of multicellular eukaryotes (Ariyoshi and Schwabe, 2003). Like FCA, FPA functions to control the site of RNA 3’ end formation and, like FCA, this was first discovered because FPA appears to autoregulate its expression by alternative polyadenylation. FPA promotes the selection of promoter-proximal poly(A) sites within the first intron of FPA pre-mRNA (Hornyik et al., 2010b). In the absence of functional FPA, the proximal poly(A) site is not used, but when FPA is overexpressed from a transgene, it promotes selection of the proximal poly(A) site at pre-mRNA transcribed from the endogenous gene almost exclusively (Hornyik et al., 2010b). The truncated RNA appears to be non-functional and, as it lacks an in-frame stop codon, is not detectably translated (Hornyik et al., 2010b). This suggests that FPA negatively autoregulates its expression in much the same way as FCA. However, there must be some subtle differences, because although FCA expression is exquisitely sensitive to this autoregulation (Quesada et al., 2003), FPA can nevertheless be overexpressed from transgenes that contain the alternatively polyadenylated first intron (Bäurle and Dean, 2008; Hornyik et al., 2010b). The mechanism by which FPA controls 3’ end formation is not yet known, but it occurs independently of FCA and the interaction of FCA with FY (Hornyik et al., 2010b). Therefore these RNA-binding proteins function in the same genetically defined flowering pathway and both control RNA 3’ end formation, but they do so in genetically separable ways (Hornyik et al., 2010b).

The subsequent characterization of FCA, FY, FPA, and components of the polyadenylation machinery has revealed that viable mutant alleles of other highly conserved components of this complex are also late flowering, specifically because of FLC misregulation. Qingshun Quinn Li took a reverse genetics approach to characterize the mRNA cleavage and polyadenylation factors of A. thaliana and discovered that mutants disrupted in a homologue of the Pcf11 protein called PCF11p-Similar protein 4 (Pfs4) flower late (Xing et al., 2008). Pcf11 is a highly conserved core component of the cleavage and polyadenylation machinery that can also function directly in transcription termination (Zhang et al., 2005; Bin Tian and Manley, 2013). Viable mutant alleles of two other core conserved components of the cleavage and polyadenylation machinery: cleavage stimulation factors CSTF77 and CSTF64 were identified in a mutant screen designed to identify factors required for FCA function. In cstf77 and cstf64 mutants, flowering was delayed and levels of FLC were elevated (Liu et al., 2010).

The late-flowering phenotype of factors that function in RNA 3’ end formation does not result from some overall slowing of development resulting from the misregulation of multiple genes. Instead, the late flowering of these mutants can be explained by the specific upregulation of FLC, as double mutants that lack functional FLC and are defective in either Pcf4 (Xing et al., 2008), FCA (Michaels and Amasino, 2001), FY (Henderson et al., 2005), or FPA (Michaels and Amasino, 2001) are not late flowering. Together, these findings reinforce the connection between RNA 3’ end formation and the control of flowering through the specific control of FLC transcription.

Although upstream input signals into the autonomous pathway are not known and many components are highly conserved core factors of the cleavage and polyadenylation machinery, they can nevertheless be subject to regulation. For example, the autoregulation of FCA provides exquisite control of FCA activity, but there is a temporal and spatial bias to this control (Macknight et al., 2002). What mediates this cell-specific control of FCA alternative polyadenylation is not known, but it can affect the local activity of FCA and flowering time (Macknight et al., 2002). In addition, sequences in the 5’ untranslated region of FCA are required for non-canonical translation initiation, which potentially allows tighter control over FCA expression (Simpson et al., 2003).
Other RNA-processing proteins control FLC expression

Late-flowering mutants

Since the characterization of Koornneef’s first autonomous pathway mutants, other late-flowering mutants have been discovered that share the same phenotypes, and these extend beyond proteins known to function in RNA 3′ end formation. These include other RNA-binding proteins such as FLK, which comprises three KH-type RNA-binding domains (Lim et al., 2004; Mockler et al., 2004), but the function of which is unknown, and other RNA-binding proteins that have been shown to affect pre-mRNA splicing. For example, A. thaliana glycine-rich RNA-binding protein 7 (AtGRP7) and AtGRP8 can autoregulate and cross-regulate their expression through regulation of alternative splicing and affect flowering time through FLC regulation (Streitner et al., 2008). Additionally, two splicing factors, Serine/arginine rich protein 45 (SR45; Ali et al., 2007) and pre-mRNA processing 39-1 (PRP39-1; Wang et al., 2007) regulate FLC.

Early flowering mutants

The late-flowering phenotype of FRI or autonomous pathway mutant backgrounds can be suppressed by mutation of a number of factors required for FLC expression, and several of these are RNA-processing factors. For example, mutations in the large or small subunits of the nuclear mRNA cap-binding protein complex, CBP80 (ABA hypersensitive 1, ABA1) and CBP20, compromise FLC expression (Bezerra et al., 2004; Geraldo et al., 2009). This may be because of defective splicing of FLC intron 1 (Kuhn et al., 2007). The RNA-binding protein SERRATE (SE) is closely associated with the cap-binding complex and has overlapping targets in RNA processing (Laubinger et al., 2008). Accordingly, se mutants also flower early because of reduced levels of FLC.

Interestingly, PEPPER (PEP), which, like FLK, has three KH domains (Ripoll et al., 2006), also regulates FLC (Ripoll et al., 2009). However, PEP antagonizes FLK function, because it promotes FLC expression, as evidenced by the fact that the pep mutation partially suppresses the late-flowering phenotype of fkl (Ripoll et al., 2009).

Mutations that disrupt components of the RNA polymerase (Pol) II-associated factor complex (PAF1c) are also early flowering because they are required for FLC expression (Yu and Michaels, 2010). The PAF1c complex appears to play multiple roles in transcription elongation, including the recruitment of factors that modify chromatin and execute RNA processing. The disruption of PAF1c complexes in yeast ultimately results in defective RNA termination (Sheldon et al., 2005), and in human cells the PAF1c component CDC73 directly interacts with core components of the cleavage and polyadenylation machinery (Rozenblatt-Rosen et al., 2009). Furthermore, transcription-stimulated coupling of 3′ RNA processing requires PAF1c (Nagaike et al., 2011). Thus, while A. thaliana mutants defective in PAF1c complexes exhibit reductions in H3K4me3 (Yu and Michaels, 2010) consistent with reduced transcription, an uncharacterized feature of their impact may be associations with RNA 3′ end processing.

The FLC locus is associated with a complex set of sense and antisense transcripts.

There is little evidence that FLC pre-mRNA is alternatively spliced or alternatively polyadenylated in mutant backgrounds defective in the RNA-processing factors of the autonomous pathway (Duc et al., 2013). Single-molecule direct RNA sequencing suggests that there is no major change in the patterns of FLC polyadenylation between WT and fpa mutants, for example (Duc et al., 2013). Nor does the expression of FLC change in exosome subunit knockout backgrounds, implying that the change in FLC pre-mRNA levels between WT and autonomous pathway mutants is not explained by exosome-mediated RNA turnover (Chekanova et al., 2007). Instead, available evidence suggests that the misregulation of FLC in autonomous pathway mutants takes place at the transcriptional level (Liu et al., 2010). This means that these RNA-processing proteins function indirectly to regulate an intermediate(s) that controls FLC transcription.

Although sense-strand FLC pre-mRNA appears not to be subject to alternative processing, the FLC locus is replete with non-coding antisense RNAs that are (Fig. 1). Antisense RNAs are transcribed from a region downstream of the 3′ untranslated region of FLC (Liu et al., 2007; Swiezewski et al., 2007). Although antisense RNAs can initiate from nucleosome-free regions of poly(A) sites (Jacquier, 2009), there is evidence that this region of the FLC locus can function as an authentic promoter for antisense transcription independent of FLC sense-strand transcription (Sun et al., 2013). The transcription start site(s) of these antisense RNAs is ill-defined, but antisense RNAs that are both alternatively spliced and alternatively polyadenylated can be detected (Hornyik et al., 2010b). The intron splice sites of FLC antisense RNAs match U2-type consensus sequences and are distinct in position from those that comprise sense-strand FLC pre-mRNA (Hornyik et al., 2010b). Cleavage and polyadenylation take place at promoter-proximal sites that are coded on the strand opposite to that coding for FLC intron 6 and distal sites that are coded on the opposite strand of a position upstream of the FLC transcription start site (Fig. 1).

In autonomous pathway mutants, more antisense RNA polyadenylated at the distal poly(A) sites is detected (Hornyik et al., 2010b). This finding is confirmed by single-molecule direct RNA sequencing, which reveals a statistically significant increase in read counts in distally polyadenylated antisense RNAs in fpa mutants (Duc et al., 2013). Single-molecule direct RNA sequencing has the advantage that it simultaneously detects transcripts cleaved and polyadenylated at multiple sites, typical of the heterogeneous pattern of RNA 3′ end formation in plants (Sherstnev et al., 2012). In other words, in genotypes with increased FLC sense-strand expression, increased antisense expression is detected too. The relative
levels of polyadenylated transcripts detected are quite different, however, with antisense expression being approximately two orders of magnitude lower than sense-strand expression (Duc et al., 2013).

There is not yet a unified view of events affecting proximal polyadenylation of *FLC* antisense RNAs. Using oligo(dT)-primed reverse transcription and a primer combination first used by Liu et al. (2007), no apparent reduction in the proximally polyadenylated antisense RNAs was detected in *fpa* mutants (Hornyik et al., 2010b). However, there are weaknesses in this approach because proximally cleaved polyadenylation sites are not measured directly and artefactual signals resulting from internal priming may be measured too (Sherstnev et al., 2012). Using a primer combination that detects polyadenylation at a single site, Liu et al. (2010) interpreted quantitative RT-PCR data to suggest that proximal poly(A) site selection in antisense transcripts is reduced in autonomous pathway mutants. However, this approach is also limited by the fact that only a single cleavage site is detected, which may not be the preferred cleavage site or be representative of the heterogeneous processing of the 3′ end of these RNAs. Indeed, a reduction in proximally polyadenylated *FLC* antisense RNAs was not clearly substantiated by single-molecule direct RNA sequencing of *fpa* mutants (Duc et al., 2013). The abundance of such RNAs is low and it is not clear how accessible all RNAs are to this technology, so this finding would benefit from further analysis. However, although the ratio of cleavage and polyadenylation at proximal and distal sites is certainly different, and the amount of cleavage and polyadenylation at distal sites increases, the amount of cleavage and polyadenylation detected at the proximal sites of antisense RNAs may not change in autonomous pathway mutants.

Another feature adding to the apparent complexity at the *FLC* locus is that R-loops, formed between nascent RNA invading the complementary strand of DNA as it is transcribed, are detected in a region that overlaps the proximally polyadenylated antisense RNAs (Sun et al., 2013) (Fig. 1). In other words, some fraction of the antisense RNAs detected in this region are engaged in R-loop formation. In *A. thaliana*, the AtNDX homeodomain protein binds and stabilizes these R-loops at *FLC* (Sun et al., 2013). In the absence of AtNDX, *FLC* levels increase and antisense RNA expression increases. Double mutants between autonomous pathway mutants and AtNDX mutants do not result in further increases in *FLC* expression, suggesting that the impact of R-loops on *FLC* expression ultimately overlaps that of the autonomous pathway (Sun et al., 2013). R-loops are associated with transcription start sites and transcription termination sites in many human genes (Ginno et al., 2013). R-loops associated with
transcription termination regions can promote termination by pausing Pol II elongation (Skourtì-Stathaki et al., 2011). It is not yet known whether antisense RNAs engaged in R-loops can be detected by single-molecule direct RNA sequencing, either because such RNAs are not polyadenylated or because their association with DNA prevents their purification and detection with other RNAs (Duc et al., 2013).

The cast of transcripts at the FLC locus includes another non-coding RNA known as COLDAIR (COLD ASSISTED INTRONIC NONCODING RNA) (Fig. 1). This RNA is transcribed in a sense direction from a promoter within the first intron of FLC. Unusually, COLDAIR appears to be transcribed by Pol II and capped, but it is not spliced or polyadenylated (Heo and Sung, 2011). It has not yet been reported whether the expression of this RNA is altered in autonomous pathway mutant backgrounds.

How might 3′ end formation of non-coding antisense RNAs regulate FLC?

It is possible that the processing of antisense RNA controls FLC expression. FCA and FPA localize to FLC chromatin, suggesting that they may directly affect RNA 3′ processing at this locus (Liu et al., 2007; Hornyik et al., 2010b). There is well-established interplay between splicing and polyadenylation, with splicing of terminal exons stimulating cleavage and polyadenylation (Niwa et al., 1990). Likewise, interactions between Pol II elongation, chromatin modification, and splicing are increasingly well documented (Kornblihtt et al., 2009). Therefore, one can readily conceive that the diverse components of the autonomous pathway, which include factors that influence chromatin modification, splicing, polyadenylation, termination, and other RNA-processing factors that affect FLC expression, might do so by influencing alternative patterns of transcription termination of antisense RNAs at the FLC locus (Hornyik et al., 2010a).

It has been suggested that FCA and FPA play widespread roles in RNA-mediated chromatin silencing (Bäurle et al., 2007), and this has informed some interpretations of how these proteins might function in flowering time control. However, the idea that these proteins function in RNA-mediated chromatin silencing appears to be mistaken and can instead be explained by the role these proteins play in RNA 3′ end formation. Bäurle et al. (2007) reported that the AtSN1 SINE retroelement was upregulated in fpa mutants. AtSN1 is a target of the RNA-directed DNA methylation pathway and thus typically silenced by DNA methylation. However, DNA methylation at AtSN1 is unchanged in fpa mutants (Bäurle et al., 2007). Hornyik et al. (2010b) explained this finding through FPA’s role in controlling RNA 3′ end formation, revealing that, in the absence of FPA, defective termination at an upstream Pol II-transcribed gene results in read-through transcription into intergenic space and over the AtSN1 locus. Subsequent independent genome-wide studies also fail to support a widespread role for these proteins in RNA silencing. First, analysis of DNA methylation in 82 silencing mutant backgrounds, including fca and fpa, revealed that changes in DNA methylation patterns were not shared between fca fpa and mutants of the RNA-directed DNA methylation pathway (Stroud et al., 2013). Secondly, single-molecule direct sequencing of RNA from fpa mutants failed to find widespread misregulation of targets of the RNA-directed DNA methylation pathway (Duc et al., 2013). Therefore, RNA 3′ end formation and its intimate connection to transcription termination appear to account for the way these proteins function.

How these RNA-processing proteins actually control FLC transcription remains unclear. It has also been suggested that targeted 3′ end processing of the proximal antisense RNA triggers silencing of FLC chromatin (Liu et al., 2010). This is based on the observation that both FPA and FCA can be detected in close proximity to this cleavage event in chromatin immunoprecipitation analyses (Liu et al., 2007; Hornyik et al., 2010b) and that the full activity of FCA and FPA in controlling FLC expression appears to require the histone demethylase FLD (Liu et al., 2007). Consistent with this idea, increased levels of H3K4me2 are found at FLC chromatin downstream of the proximal poly(A) sites of antisense RNAs in fld mutants (Liu et al., 2007). However, other data do not substantiate this idea. For example, fld mutants are not completely epistatic to overexpressed FCA or FPA, so it is not absolutely clear that histone demethylation is consequently downstream of their role in RNA 3′ end formation (Liu et al., 2007). Secondly, when combined with mutations in the PAF1c component CDC73, the response of fca and fld mutants can be separated: although the late-flowering phenotype of fca mutants is suppressed in cdc73 mutant backgrounds, cdc73 fld mutants still flower late (Yu and Michaels, 2010). This finding is inconsistent with the idea that FLD functions obligatorily downstream of FCA. Thirdly, using single-molecule direct RNA sequencing to quantify shifts in poly(A) site selection, Duc et al. (2013) found no clear change in the usage of proximal poly(A) sites in fpa mutants, although increased amounts of transcripts polyadenylated at promoter distal sites were clearly detected. If cleavage at proximal poly(A) sites in antisense RNA triggers silencing of FLC chromatin, it is unclear why silencing would be lost when proximal cleavage continues to occur in autonomous pathway mutants (Duc et al., 2013). Finally, none of the FLC single-nucleotide polymorphisms (SNPs) in natural Arabidopsis thaliana accessions so far associated with flowering time and antisense expression levels map in the proximity of proximally polyadenylated antisense RNAs (Coustham et al., 2012). Instead, SNP-121, which is correlated with flowering time (in the absence of vernalization) and antisense RNA expression levels, maps to the FLC promoter in a region corresponding to sequence encoding a distally polyadenylated antisense RNA poly(A) signal (Coustham et al., 2012; Duc et al., 2013).

If antisense RNA processing does indeed regulate FLC transcription, it might be that defective termination at proximal poly(A) sites results in read-through RNAs traversing the FLC promoter and stimulating sense-strand expression. It is conceivable that sense and antisense transcription are intertwined, such that sense-strand transcription enhances antisense transcription, enhances sense-strand
transcription’ in a self-reinforcing feedback loop, similar to that described recently for yeast Arg1 (Crisucci and Arndt, 2012). Nucleosome-free regions formed at promoter and terminator regions may facilitate transcription on the other strand (Jaucquier, 2009). In addition, if gene loops form between the promoter and terminator regions of FLC, facilitating local recycling of Pol II and transcription memory (O’Sullivan et al., 2004; Ansari and Hampsey, 2005; Tan-Wong et al., 2009, 2012), they would also bring together the promoter and terminator region of antisense RNAs. Notably, gene loops require cleavage and polyadenylation factors for their formation (Ansari and Hampsey, 2005). Transcription of non-coding antisense RNAs through promoter regions can alter nucleosomes and thus change sense-strand transcription. In many cases, sense-strand transcription is compromised (Castelnuovo et al., 2013), but there are examples where sense-strand expression is increased and thus, like FLC, where sense and antisense expression positively correlate (Uhler et al., 2007; Crisucci and Arndt, 2012). Indeed transcriptome-wide sequencing of fpa mutants revealed this positive correlation between sense and antisense expression not only at FLC but also at a small number of other genes (Duc et al., 2013). Thus the FLC locus resembles yeast loci in which mechanisms are in place to promote early termination of non-coding antisense RNAs to prevent them traversing sense-strand promoter regions where they otherwise enable altered transcription responses (Crisucci and Arndt, 2012; Castelnuovo et al., 2013).

Current limitations to the concept of antisense RNA-based regulation of FLC

The proposed role of RNA-processing factors in regulating FLC transcription through their action on terminating antisense RNA transcription currently depends on data that derive from correlation. In autonomous pathway mutants, there is increased read-through of distally polyadenylated antisense RNAs traversing the FLC promoter, and there is an increase in FLC sense-strand expression. However, it is not clear whether changes in polyadenylated antisense RNAs cause changes in FLC transcription or are a consequence of altered FLC transcription instead. Indeed, it is formally possible that the changes in FLC sense-strand transcription and antisense RNA polyadenylation are merely coincidental. The genetic separation of these events is not entirely straightforward because the promoter, coding, and poly(A) signal sequences of these RNAs are effectively embedded within another, but this will be crucial for separating cause from effect. So far, analyses have been based upon RNA purified from whole seedlings, but it will be important to define what is happening at individual loci in individual cells, because sense and antisense expression at the same locus may be separated by time or occur mutually exclusively in different cells, as reported recently for non-coding antisense RNAs in yeast (Castelnuovo et al., 2013).

If not antisense RNA processing, then what?

If changes in antisense RNA processing are a coincidence of changes in FLC transcription, the mechanism by which the RNA-processing components of the autonomous pathway might control the transcription of FLC remains unclear. Multiple changes in gene expression occur in fca and fpa mutants (Hornyik et al., 2010b; Sonmez et al., 2011; Duc et al., 2013; Lyons et al., 2013). Recent single-molecule direct RNA sequencing analysis quantified shifts in expression and RNA 3’ end formation between wild-type and fpa mutant transcriptomes (Duc et al., 2013). Although limited by sequencing depth, sequence read length, and the detection only of polyadenylated RNAs, a range of consequences of disrupting FPA function were uncovered using this approach (Duc et al., 2013; Lyons et al., 2013). For example, specific cases of intragenic alternative polyadenylation and defective 3’ end formation that resulted in intergenic transcription were found. The consequences of read-through transcription included the formation of novel antisense RNA transcripts, transcription through the promoters of neighbouring genes and chimaeric RNAs formed from the exons of two otherwise separate and well-characterized genes that were also associated with the silencing of native expression of the downstream gene (Duc et al., 2013). Intergenic read-through in fpa mutants is not necessarily benign: read-through at the single-exon ERF4 gene, for example, resulted in cleavage and polyadenylation occurring at intergenic sites accompanied by a cryptic splicing event that altered the coding sequence of ERF4 mRNA and the function of the ERF4 protein (Lyons et al., 2013). A subset of intragenic heterochromatin sequences that depend upon increase Bonsai Methylation 2 (IBM2) for their transcription (Saze et al., 2013) were readily transcribed in fpa mutants (Duc et al., 2013). Consequently, the span of disruptions on gene expression in fpa mutants is wide, and, until proven otherwise, it remains formally possible that some of these events ultimately lead to the misregulation of FLC.

Distinguishing the autonomous and vernalization pathways: COOLAIR and COldaIR non-coding RNAs

The autonomous and vernalization pathways both down-regulate the transcription of FLC, but they are distinct in how they achieve this. Koornneef first showed that the vernalization response was not impaired in autonomous pathway mutants, revealing that the mechanisms involved must be different (Koornneef et al., 1991). No RNA-processing factors have yet been implicated in the vernalization response pathway; instead, silencing mediated by polycomb group complexes modifying chromatin with H3K27me3 is central to this control (Andrés and Coupland, 2012).

Different non-coding RNAs at the FLC locus have been implicated in the vernalization response. First, antisense RNAs whose expression increases in the cold (Swiezewski
et al., 2009) (and which may be related to the antisense RNAs affected in autonomous pathway mutants at ambient temperature) called COLD INDUCED LONG ANTISENSE INTRAGENIC RNA (COOLAIR) were proposed to mediate the silencing of FLC during the vernalization response (Swiezewski et al., 2009). However, the disruption of these antisense RNAs by T-DNA insertion does not compromise the vernalization response (Helliwell et al., 2011). This suggests that COOLAIR RNAs are not required for vernalization (with the caveat that the large disruption in chromatin and/or introduction of promoter and terminator sequences found in T-DNAs do not mimic the effects of COOLAIR RNAs). Because of the distinctions in the vernalization and autonomous pathway function, a role for antisense RNAs in FLC regulation through the autonomous pathway remains possible, but they appear not to be required for the vernalization response.

A second non-coding RNA at the FLC locus, COLDAIR, has been implicated in the vernalization response. COLDAIR is detectable at ambient temperature, but its expression increases in response to low temperatures (Heo and Sung, 2011). In contrast to COOLAIR, COLDAIR has been shown to be genetically required for the vernalization response and to mediate the association of polycomb group protein complexes with the FLC locus (Heo and Sung, 2011). The COLDAIR promoter is located in the vernalization responsive element present in intron 1 of FLC (Sung et al., 2006) (Fig. 1).

The wider impacts of RNA processing controlling A. thaliana flowering

One of the validations of the use of A. thaliana as a model has been the recognition of how widely discoveries first made in A. thaliana can be applied to explain facets of plant biology. The genes disrupted in the photoperiod pathway mutants identified by Koornneef are widely conserved in other plants, and proteins such as FT have been shown to promote flowering of a range of plant species (Taoka et al., 2013). In contrast, although autonomous pathway components are also widely conserved, little evidence of their role in controlling flowering outside of A. thaliana is available. This might be explained by the target of this control, FLC, not being widely conserved either (although recent analysis suggests that a sequence related to FLC is more widely conserved than had been thought; Ruelens et al., 2013) and the existence of antisense RNAs specifically at A. thaliana FLC. Consequently, the RNA-processing components of the autonomous pathway may not be general flowering regulators in other species. Consistent with this idea, autonomous pathway mutants are associated with other phenotypes. For example, FPA affects the A. thaliana immune response (Lyons et al., 2013). Perhaps the impact of alternative processing of antisense RNAs at the FLC locus associates them with a flowering function in A. thaliana that nevertheless is subject to natural occurring variation and adaptation. As a result, the general lessons learned (and that continue to be learned) from studying the autonomous pathway relate not only to flowering per se but also to how long non-coding RNAs affect gene expression, how alternative polyadenylation is regulated, and how precision in quantitative gene regulation occurs in vivo. The regulation of FLC has been studied intensely, and many mutants defective in its regulation have been isolated. And yet, we still do not understand how the transcription of FLC is controlled. In this evolving field, the complexity of gene regulation at the FLC locus still awaits clarity.

Acknowledgements

Research in Gordon Simpson’s laboratory on non-coding RNAs and alternative polyadenylation is funded by BBSRC grants BB/J00247X/1 and BB/H002286/1. GGS thanks Jane Mellor and Françoise Stutz for helpful discussions.

References


Simpson GG, Dijkwel PP, Quesada V, Henderson I, Dean C. 2003. FY is an RNA 3′ end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell 113, 777–787.


