Quantitative regulation of FLC via coordinated transcriptional initiation and elongation

Wu, Z., Ietswaart, R., Liu, F., Yang, H., Howard, M., & Dean, C. (2016). Quantitative regulation of FLC via coordinated transcriptional initiation and elongation. Proceedings of the National Academy of Sciences, 113(1), 218-223.

Published in:
Proceedings of the National Academy of Sciences

Document Version:
Early version, also known as pre-print

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Quantitative regulation of *FLC* via coordinated transcriptional initiation and elongation

Zhe Wu[1,*], Robert Ietswaart[1,2,*], Fuquan Liu[1,3], Hongchun Yang[1], Martin Howard[1,2,#] and Caroline Dean[1,#]

1. Department of Cell and Developmental Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom 2. Computational and Systems Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom 3. Present address: Institute of Global Food Security, School of Biological Sciences, Queen's University, Belfast, BT9 7BL, United Kingdom *These authors contributed equally to this work. # Co-corresponding authors.

Submitted to Proceedings of the National Academy of Sciences of the United States of America

The basis of quantitative regulation of gene expression is still poorly understood. In *Arabidopsis thaliana* quantitative variation in expression of *FLOWERING LOCUS C* (*FLC*) influences the timing of flowering. In ambient temperatures *FLC* expression is quantitatively modulated by a chromatin silencing mechanism involving alternative polyadenylation of antisense transcripts. Investigation of this mechanism unexpectedly showed that RNA Polymerase II (Pol II) occupancy changes at *FLC* did not reflect RNA fold changes. Mathematical modeling of these transcriptional dynamics predicted a tight coordination of transcriptional initiation and elongation. This prediction was validated by detailed measurements of total and chromatin-bound *FLC* intronic RNA; a methodology appropriate for analyzing elongation rate changes in a range of organisms. Transcription initiation was found to vary ~25-fold with elongation rate varying ~8-12-fold. Premature sense transcript termination contributed very little to expression differences. This quantitative variation in transcription was coincident with variation in H3K36me3 and H3K4me2 over the *FLC* gene body. We propose different chromatin states coordinately influence transcriptional initiation and elongation rates and that this coordination is likely to be a general feature of quantitative gene regulation in a chromatin context.

chromatin | alternative polyadenylation | COOLAIR | autonomous pathway | FCA

Introduction

The influence of chromatin on transcription and co-transcriptional processing is of central importance in the regulation of gene expression (1, 2). An intensively studied example where the local chromatin state is considered to influence transcription in *Arabidopsis* is *FLOWERING LOCUS C (FLC)*. *FLC* encodes a MADS-box transcription factor and acts as a floral repressor (3, 4). *FLC* expression is tuned by different genetic pathways: FRIGIDA activates *FLC* expression through a mechanism requiring Trithorax homologues, Paf1C and SDG8, an H3K36 methyltransferase (5). *FLC* expression is repressed by the autonomous pathway and vernalization (5). Both these repressive pathways involve a group of antisense long non-coding transcripts collectively termed *COOLAIR*, which initiate immediately downstream of the poly A site at the 3' end of *FLC*. These antisense transcripts terminate at either proximal sites internal to the *FLC* gene, or distal sites within the *FLC* promoter (6, 7). Mutation of autonomous pathway components, including the RNA binding proteins FCA and FPA and the conserved components of the 3' processing complex FY, Cstf64 and Cstf77 leads to relative reduction in use of the proximal polyadenylation sites and increased *FLC* sense expression (reviewed in (8)). FCA localizes to *FLC* chromatin near the proximal poly A sites (9), and this together with the fact that PRP8 and CDKC2 (P-TEFb component), identified in FCA suppressor screens (10, 11), both require *COOLAIR* to repress *FLC,* supports the idea that promotion of proximal polyadenylation of *COOLAIR* is directly linked to reduced *FLC* expression. FLD, an H3K4me2 demethylase, also functions in this mechanism and *fld* is the most effective suppressor of FCA function at *FLC* (9). FLD modulates H3K4me2 levels in the gene body of *FLC,* however, how FCA functions with FLD to achieve *FLC* repression remains to be fully elucidated.

Here, we investigate how FCA and FLD transcriptionally repress *FLC* through analysis of Pol II occupancy. We use these data together with RNA measurements to parameterize an analytic mathematical model of *FLC* transcription. Model predictions are then tested through detailed measurements of intronic total and chromatin-bound RNA levels. This methodology is very appropriate for evaluating elongation rate changes in whole organisms where pulse-chase experiments are technically unfeasible. At *FLC*, we find that both FCA and FLD-mediated repression occurs not only through reduced transcription initiation, but also through a coordinately reduced Pol II elongation rate. We propose that chromatin modifications at *FLC* induced by FCA and FLD, influenced by the antisense transcript processing, coordinately change initiation and elongation to quantitatively regulate the transcriptional output of the locus.

Results

RNA fold changes do not reflect Pol II occupancy changes

Measurement of steady state spliced *FLC* and unspliced *FLC* RNA showed an increase in expression of ~20-25 fold between Col and *fca-9* and *fld-4* (Fig. 1A). We reasoned that if this was caused by a 25x change in transcription initiation a 25x increase in Pol II levels would be found at *FLC*, assuming transcript half-lives, splicing/3' processing efficiency, Pol II processivity and

Significance

The textbook view of how transcription is quantitatively regulated is through changes in transcription initiation. However, the arrangement of DNA in chromatin in eukaryotes and the frequent occurrence of non-coding transcripts add to the complexity of transcriptional regulation. Here, we explore the quantitative transcriptional regulation of *FLC*, a gene important for developmental timing in Arabidopsis. *FLC* expression correlates with altered antisense transcript processing and different chromatin states. Through experiments and mathematical modeling we discover that transcription initiation and elongation are tightly coordinated and both are influenced by the chromatin state at the locus. Modulation of the chromatin environment by non-coding transcripts to coordinately influence transcription initiation and elongation could be a general mechanism to regulate quantitative transcriptional output.

Reserved for Publication Footnotes

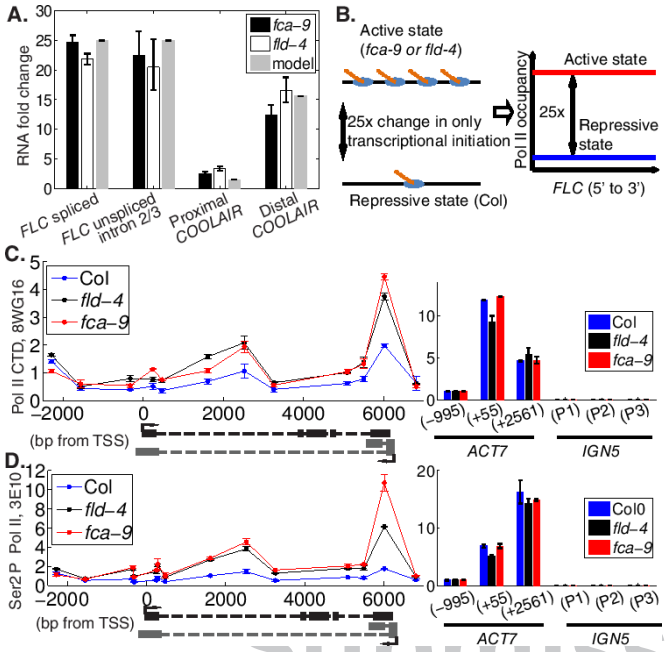

Fig. 1. Large increases in RNA are associated with small changes in Pol II occupancy (A)RNA fold up-regulation in *fca-9* and *fld-4* mutants compared to Col: spliced and unspliced *FLC* (∼25x), proximal (∼2x) and distal *COOLAIR* (∼13x). The model values are the fits to the experimental data. Experimental values are mean ± s.e.m. from 3 to 6 independent samples. (B) Schematic illustration of a scenario where transcription initiation is the only difference between Col and *fca-9*, so that a 25x fold change in Pol II occupancy should be observed as illustrated on the right. (C-D) ChIP experiments assaying Pol II occupancy across *FLC* using the antibodies anti CTD 8WG16 (C) and anti Ser^2 P CTD 3E10 (D). The bar charts at the bottom indicate Pol II levels at various control genes. Three overlapping primer pairs are used to measure *IGN5* expression (P1-P3). Values are mean ± s.e.m. from 2 independent samples, with data presented as the ratio of Pol II at *FLC* / input at *FLC* to Pol II at *ACT7* (-995) / input at *ACT7* (-995).

elongation rates are unaffected in *fca-9* and *fld-4* (Fig. 1B). However, both total Pol II and productively elongating Pol II (Ser2-P) showed relatively small changes (2-3x) across *FLC* in the different genotypes (Fig. 1C,D; Fig. S1A,B). We ruled out a number of technical issues with Pol II ChIP that could have led to an underestimation of Pol II occupancy. First, measurements on a highly expressed gene (*ACT7*) and a Pol IV/V transcribed region (*IGN5*) showed that a wide dynamic range (>1000x by comparing levels at *ACT7* to *IGN5*) could be detected in the Pol II ChIP assay (Fig. 1C,D). Pol II levels at *FLC* were well above background at *IGN5* (Fig. 1C,D; Fig. S1). Second, specific dilutions of *FLC* chromatin, without changing the overall amount of chromatin, showed rough linearity between the Pol II ChIP signal and the Pol II concentration at *FLC* (Fig. S2). Third, cell-specific *FLC* expression variation is also highly unlikely to underlie this difference in RNA and Pol II up-regulation, as both assays use whole plant seedlings and thus reflect population averages. Based on these observations, we conclude that FCA/FLD-mediated changes in *FLC* transcription are unlikely to occur solely through changes in transcription initiation.

***FLC* transcriptional dynamics can be explained by coordination of initiation and elongation**

To further understand how FCA and FLD-mediated *FLC* repression occurs at a transcriptional level, we developed an analytical mathematical model of the transcriptional dynamics at *FLC* by incorporating sense *FLC* and *COOLAIR* initiation, elongation and termination (Fig. 2A; see Supporting Information for complete description). The experimental data described above were used as model inputs. This strategy enabled us to

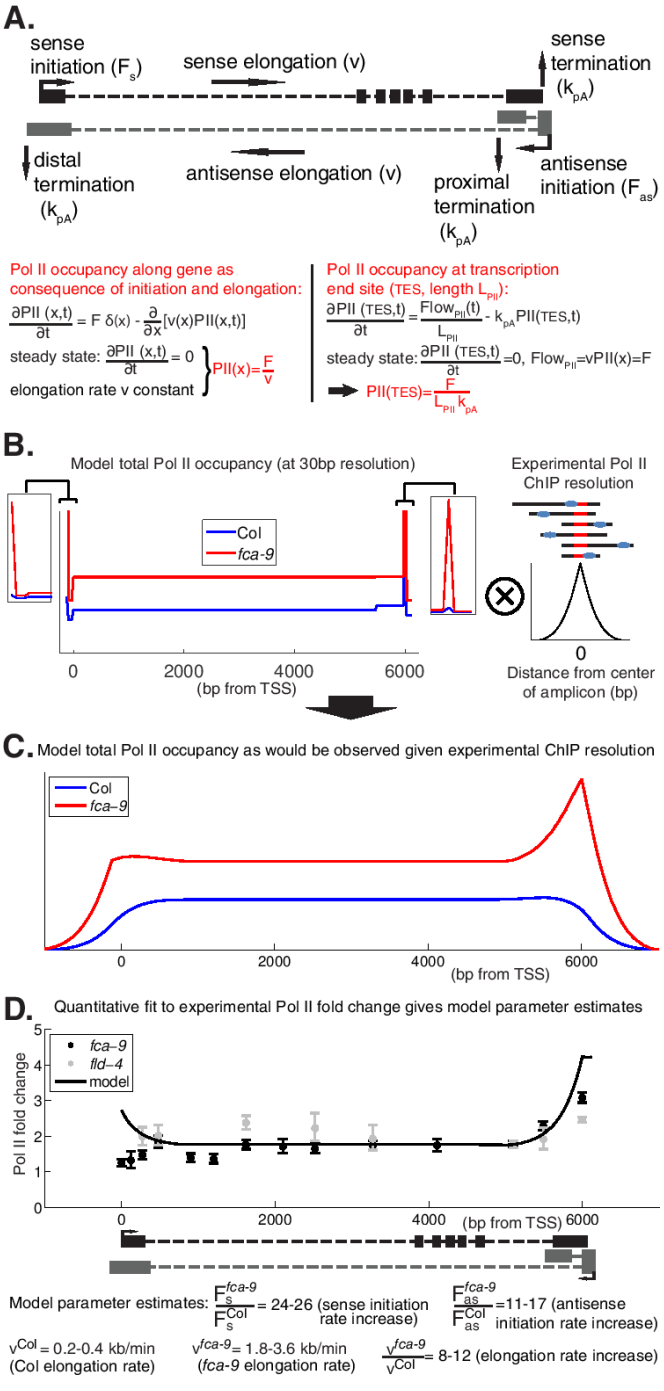

Fig. 2. Small changes in Pol II occupancy can be explained by coordinated changes in transcription initiation and elongation(A)Schematic of *FLC* locus andoutline of the mathematical model for *FLC* transcription (details in Supporting Information). Black boxes indicate sense exons; grey boxes: proximal (upper) and distal (lower) antisense exons.(B) Total (sum of sense and antisense) model Pol II levels in Col and *fca-9* across *FLC*. The *fld-4* mutant model results are identical to *fca-9*. Shown on the right is a schematic of the convolution process with experimental Pol II ChIP fragment size distribution (shown in Fig. S3).(C) Total Pol II levels in Col and *fca-9* across *FLC* from the model convolved with experimental Pol II ChIP fragment size distribution.(D) Experimental and model Pol II fold up-regulation. Experimental values are mean ± s.e.m. from 2 to 5 independent samples, including data shown in Fig. 1C,D and Fig. S1. Model fold changes are ratio of profiles shown in (C).

assign parameter values for key processes during transcription (e.g. initiation and elongation). Pol II levels reflect a density that

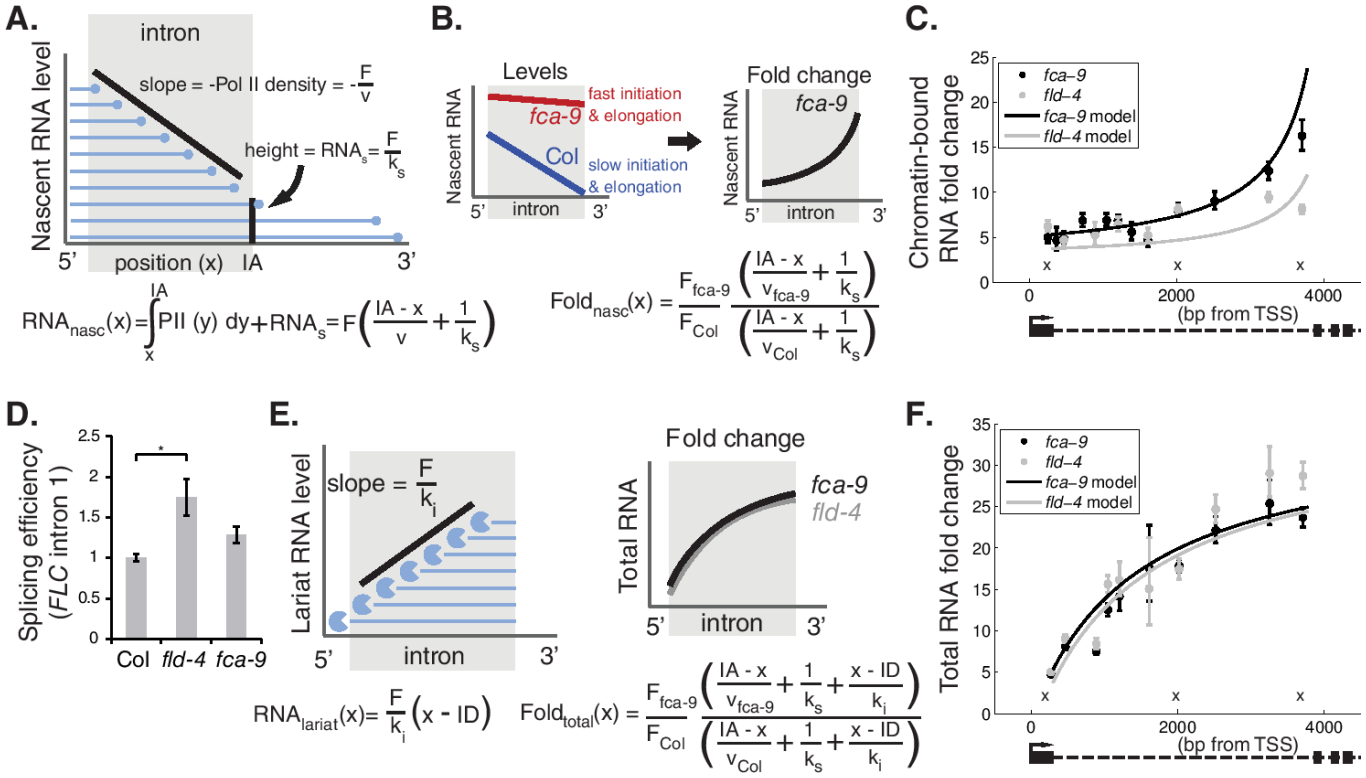

Fig. 3. Combination of increased initiation and elongation, with co-transcriptional splicing and lariat degradation, leads to distinct RNA profiles along *FLC* intron1 (A) Schematic indicating intronic nascent RNA, RNA_{nasc} (blue lines), arising from Pol II (blue circles) elongating through the intron and from unspliced RNA_s with full-length intron. Once Pol II has passed the intron acceptor site (IA), splicing can occur. Initiation, elongation and splicing rates are respectively F, v and k_s. Analytic expression for RNA_{nasc} shown below. (B) Schematic (left panel) indicating model profiles of nascent RNA along *FLC* intron1 in *fca-9* and Col. Between *fca-9* and Col, F and v are coordinately increased, but with the same k_s. This generates a characteristic pattern of intronic nascent RNA fold changes between *fca-9* and Col (right panel) with analytic expression shown. (C) Modeled and experimentally measured chromatin-bound RNA fold changes along *FLC* intron1. The lower increase towards the 3' end in *fld-4* is due to increased splicing rate as shown experimentally in (D). Crosses indicate positions where data are from 3 different, overlapping primer sets that each show similar results (Fig. S4). (D) Estimate of *FLC* intron1 splicing efficiency (intron cleavage rate) in *fld-4* and *fca-9*, normalized to the level in Col. Values are mean ± s.e.m. from 3 independent samples. Asterisks indicate statistical significance: for all the figures in this study, *p < 0.05, **p<0.01, ***p<0.001, two-sided unpaired t-test unless specified otherwise. (E) Schematic showing effect of 5' to 3' intronic RNA degradation on lariat RNA levels (RNA_{lariat}). Full-length lariat RNA results from splicing and is degraded with rate k_i, ID: intron donor site. These degradation intermediates, together with the nascent RNA described in (A), make up total intronic RNA. Fold up-regulation then generates the characteristic profiles shown. Analytic expressions for RNA_{lariat} and total intronic RNA fold changes shown. (F) Modeled and experimentally measured total RNA fold changes along *FLC* intron1. (C and F) Experimental values are mean ± s.e.m. from at least 3 independent samples. Absolute levels are shown in Fig. S4.

can be described mathematically as a ratio of the initiation rate (F) over the elongation rate (v) (12). Since our ChIP signal is not strand specific, we summed the sense and antisense model Pol II levels to generate a model total Pol II profile along *FLC* (Fig. 2B). The small increase of Pol II ChIP signal in the transcriptionally active *fca-9* and *fld-4* mutants (Fig. 1C,D; Fig. S1) is explained by the model through a coordinated increase in initiation and elongation rates (Fig. 2B,C). The model also reproduced the *FLC* spliced, unspliced and *COOLAIR* fold up-regulation in *fca-9* and *fld-4* (Fig. 1A), where a 25x fold increase in sense Pol II initiation required an 8-12x fold faster rate of elongation to quantitatively fit the Pol II occupancy increase (Fig. 2D). Elevated Pol II levels at the 3' of *FLC* resulted from sense termination and proximal antisense transcription (Fig. 2A-D). Our model does not take into account transcriptional interference (TI) between sense *FLC* and *COOLAIR* (see Discussion). Using an experimentally determined value for the termination rate 1/50 s^{-1} (13), absolute elongation rates could be inferred from the model, yielding 0.2-0.4 kb/min (Col) and 1.8-3.6 kb/min (*fca-9* and *fld-4*). These correspond well to values found in other organisms (14-17). The excellent fit of the experimental data strongly supports a model where *FLC* transcriptional dynamics are governed by coordinated changes in initiation and elongation.

Co-transcriptional splicing, combined with coordinated initiation and elongation, generate distinctive patterns of RNA up-regulation along *FLC* intron1

We next tested the predicted coordinate increase in initiation and elongation rates experimentally. Measurement of elongation rates on a subset of highly expressed, long mammalian genes (>50 kb) has been achieved using GRO-seq (14). This technique involves inhibition of elongation and then release and relies on rapid removal of an inhibitor that is difficult in whole organisms (15, 16). We tried an alternative approach via generation of an *FLC-MS2* fusion (13), but this was not expressed at a sufficiently high level to be useful. To overcome these limitations, we used our theoretical model to make specific predictions with regards to intronic *FLC* RNA production, which we then tested experimentally. If introns are spliced co-transcriptionally once Pol II has reached the 3' end of the intron, then nascent RNA from the 5' end of the intron resides on the chromatin longer than that from the 3' end. This generates a nascent RNA profile along an intron with declining levels from the 5' to 3' end (17, 18). An analytic mathematical analysis (Fig 3A, Supporting Information) predicts that the ratio of Pol II initiation (F) over the elongation rate (v) determines the slope of the nascent intronic RNA levels between the 5' to 3' ends, whilst the initiation rate over the splicing rate (k_s) determines the levels of completely transcribed, unspliced

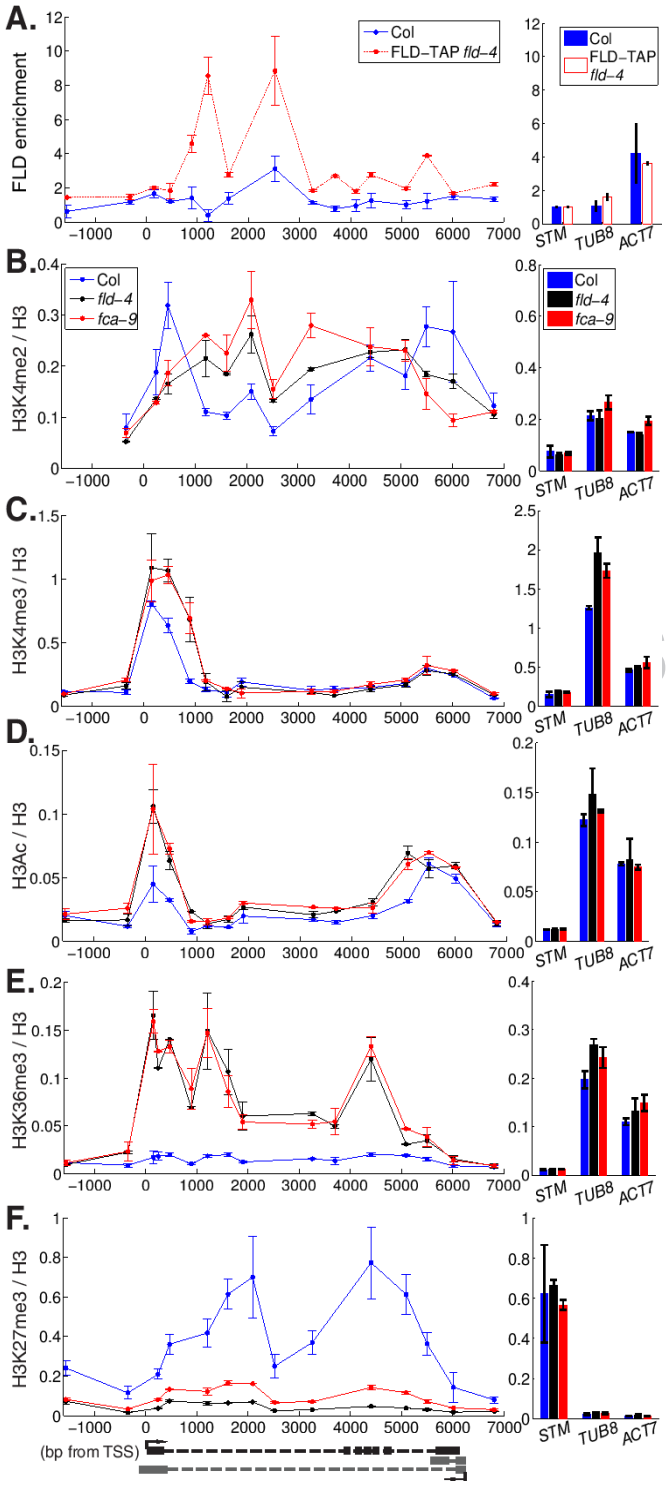

Fig. 4. FLD enrichment at the *FLC* locus is associated with changed histone modifications (A) FLD-TAP ChIP enrichment across *FLC* in Col and FLD-TAP/*fld-4*. Values are mean ± s.e.m. from 2 independent samples, with data presented as enrichment at *FLC* relative to enrichment at *STM*. (B-F) ChIP across *FLC* in Col, *fca-9* and *fld-4* measuring H3K4me2 (B), H3K4me3 (C), H3Ac (D), H3K36me3 (E), H3K27me3 (F). Values are mean ± s.e.m. from 2 independent samples, with data normalized to H3. Values at the control genes *STM*, *ACT7* and *TUB8* are shown on the right. H3/input values can be found in Fig. S7.

introns (Fig. 3A). This analysis indicates that nascent RNA levels close to the intron 3' end will be mostly determined by the ratio

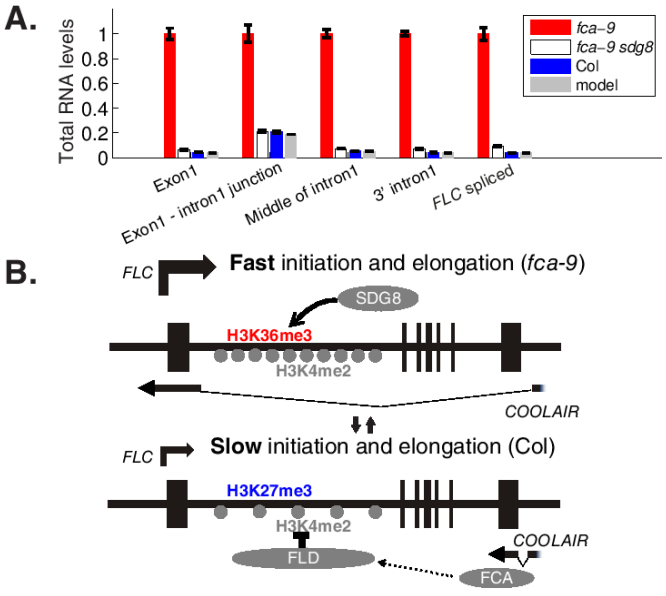

Fig. 5. Coordination of initiation and elongation at *FLC* in the H3K36 methyltransferasedeficient *sdg8* mutant. (A) Total RNA levels along *FLC* intron1. Model as described in Fig. 2. All values are relative to *fca-9*. Experimental values are mean ± s.e.m. from 3 independent samples, and are averaged from overlapping primer sets (Fig. S8). (B) Working model of how *FLC* expression is quantitatively regulated through coordination of transcription initiation and elongation. In the absense of FCA/FLD, H3K36me3 is increased at *FLC* through SDG8 function and this promotes fast transcription initiation and elongation. In presense of FCA/FLD, antisense processing triggers a reduction of H3K4me2, loss of H3K36me3 and an increase in H3K27me3, which reduces transcription initiation and slows elongation.

of the initiation rate to the splicing rate, and independent of the elongation rate. Away from the 3' end of the intron, transcripts emerging from Pol II still transcribing the intron will also contribute to nascent RNA levels, and hence the ratio of the initiation rate to the elongation rate will also be important (Fig. 3A). Taking into account both increased initiation and elongation rates in the *fca-9* mutant compared to Col (Fig. 3B), this analysis enabled us to predict a spatially varying fold up-regulation of nascent RNA along *FLC* intron1 (Fig. 3B).

We tested this key model prediction by measuring the chromatin-bound RNA profile at *FLC* (Fig. 3C; Fig. S4). Comparing *fca-9* to Col, the chromatin-bound fold up-regulation inside exon1 was much larger than at the exon1-intron1 junction (Fig. S4A,G), suggesting that splicing of intron1 does occur mostly co-transcriptionally. In the first kb of intron1, as predicted by the model, there was only a small fold increase in *fca-9* as compared to Col (Fig. 3C; Fig. S4A). This is due to the dependence on the ratios of the initiation and elongation rates and their coordinated increases in *fca-9* (Fig. 3B). By contrast, the fold up-regulation was much larger close to the intron acceptor site in *fca-9*. This is in agreement with the model, where we used the experimentally determined splicing rate of 1/100 s^{-1} (17) for both Col and *fca-9*, with other parameters determined from our prior fitting to the Pol II ChIP data (Supporting Information). Importantly, the chromatin-bound RNA profile along intron1 is not flat, which is what would be predicted without changes to the elongation rates between *fca-9* and Col.

We also fitted the model to the chromatin-bound RNA data directly using nonlinear regression (R^2=0.89, F-statistic: p=3x10^{-14}). This procedure also led to the conclusion that significant elongation rate changes (fold = 9.8±3.8 (mean±s.e.m.), p=0.03) are required to explain the profile (Supporting Information). Importantly, this method does not rely on the specific values

of splicing and elongation rates and is independent of Pol II ChIP data, and thus provides additional evidence for the elongation rate changes.

Interestingly, we observed less increase in fold up-regulation towards the 3' end of intron1 in *fld-4* as compared to *fca-9* (Fig. 3C; Fig. S4A). Given the fold change close to an intron acceptor site is more sensitive to splicing rather than elongation rate changes (Fig. 3B), we examined if a splicing rate change specific to *fld-4* could explain its differential fold up-regulation pattern from *fca-9* (Materials and Methods; Supporting Information). Indeed, we found that we could fit the *fld-4* profile in our model by incorporating a 2 fold faster splicing rate ($1/50\ s^{-1}$) in *fld-4* (Fig. 3C), whilst keeping all other parameters unchanged. We further verified this model prediction of an increased splicing rate in *fld-4* by measuring the splicing efficiency of *FLC* intron1. As predicted, the efficiency was increased 1.8 fold in *fld-4* (Fig. 3D), but not significantly altered in *fca-9* ($p=0.1$, two-sided unpaired t-test). A simple alternative model with unchanged splicing and elongation rates between Col and *fld-4* would produce a constant chromatin-bound RNA fold-change across intron1. That would be consistent with the chromatin-bound RNA data set in isolation (Fig. 3C), but implies a change in the initiation rates of ~7 fold (Supporting Information), which is inconsistent with our earlier spliced and unspliced *FLC* RNA fold changes (Fig. 1A).

To further support these conclusions we investigated the total intronic RNA profile (Fig. 3E,F; Fig. S4). Such measurements include intron lariat degradation intermediates, which are present in the total but not chromatin-bound RNA fraction (Fig. 3E)(17). Assuming that lariat degradation occurs from 5' to 3', lariat RNA at the 3' generally exists for longer than that at the 5'. This generates a lariat RNA profile with increasing levels from the 5' to 3' end (Fig. 3E). Importantly, incorporating this lariat population into the total intronic RNA fold upregulation between *fca-9* and Col, without altering the model parameterization that explained the Pol II and chromatin-bound RNA, produced a predicted profile that is qualitatively different to that found for the chromatin-bound RNA (Fig. 3B,E). This prediction was also validated experimentally (Fig. 3F). Compared to the chromatin-bound RNA profile, there was a significantly larger fold increase in the first 2kb of the total intronic RNA profile ($p = 8x10^{-7}$ and $4x10^{-7}$ for *fca-9* and *fld-4* respectively, two-sided Welch's t-test) (Fig. 3C,F; Fig. S4A,B). In the model, we could generate such a profile, by solely incorporating 5' to 3' intron lariat degradation with rate up to 1.5 bp/s (19), in line with experimentally determined intron half-lives (17). Potential additional presence of 3' to 5' degradation (19) with a rate up to 1 bp/s did not alter our conclusions (Supporting Information). The profiles for total intronic RNA look very similar between *fca-9* and *fld-4* (Fig. 3F), in contrast to the chromatin-bound data (Fig. 3C). This similarity is because the lariat RNA effectively extends the half-life of intronic RNA and therefore reduces the effect of the differential splicing rates between *fca-9* and *fld-4* (Fig. 3F). Taken together, our total and chromatin-bound intronic RNA profiles provide strong evidence that repression of *FLC* involves a coordinated change of both the initiation and elongation rates. Moreover, the methods we developed here can be used to infer elongation rate changes in whole organisms where pulse-chase experiments are not feasible.

Sense premature termination contributes little to *FLC* repression

Previous reports have linked the elongation rate to either Pol II processivity (20) or early termination (21). In these scenarios, Pol II would terminate prematurely as a result of slow elongation. Our previous analysis did not require any such premature termination. Moreover, at an intuitive level, premature termination should lead to declining levels of Pol II from 5' to 3' in the repressed case (Col) (Fig. S5A; Supporting Information). However, we found no evidence for this in our Pol II ChIP assay (Fig. 1C,D; Fig. S1) and no short transcripts had been detected by northern blot using an *FLC* intron 1 probe (22). These findings suggest that premature termination contributes little to *FLC* repression. To further confirm this conclusion, we undertook 3'RACE to map transcripts ending within the promoter-proximal region of *FLC*. We could detect polyadenylated transcripts that terminated within *FLC* intron1. These transcripts all contained *FLC* exon1 and were mostly alternatively spliced with the same donor site but with a different acceptor site, as compared to the conventional *FLC* intron1 (Fig. S5B). By monitoring the alternatively spliced intron associated with premature termination, we found these transcripts are of lower abundance than unspliced intron1 in Col, *fca-9* and *35S::FCA* (Fig. S5E). Therefore, sense premature termination occurs only occasionally at *FLC* and is not a major contributor to *FLC* repression.

Co-transcriptional decay of nascent transcripts by 5' to 3' exonucleases has also been proposed to influence transcriptional output (23, 24). In such a scenario, the degradation of RNA should also lead Pol II to terminate prematurely, and therefore to declining levels of Pol II from 5' to 3' in the repressed state (Col), which is again inconsistent with our Pol II ChIP data. In addition, we analysed *FLC* expression in mutants defective for these functions (*xrn2-1*, *xrn3-3*) (25) in *Arabidopsis* and found no increase in *FLC* nascent or fully spliced *FLC* RNA levels (Fig. S6). Therefore, such a decay pathway is unlikely to play a major role in determining the overall transcriptional dynamics at *FLC*.

FLD alters the local chromatin state to influence transcriptional output via coordinated changes in initiation and elongation

We therefore continued with our investigation of coordinated initiation and elongation rates by FCA/FLD-mediated changes in chromatin modifications. We analyzed the localization of the histone demethylase FLD at *FLC* using a complementing FLD-TAP fusion expressed from its endogenous regulatory sequences (Fig. S7A-C). FLD shows the highest enrichment at *FLC* ~1kb to 3kb downstream of the transcription start site (TSS) (Fig. 4A). This localization is consistent with the increased H3K4me2 in the *FLC* gene body (1kb to 4kb beyond the TSS) in the *fld-4* mutant (Fig. 4B). Loss of FLD, and indeed similarly FCA, resulted in changes in a number of other chromatin modifications (Fig. 4C-F). H3K4me3 and H3Ac increased around the *FLC* sense TSS (Fig. 4C,D), coincident with lower H3K4me2 in this region. The relatively small changes in H3K4me2 were correlated with much larger changes in H3K36me3 and the mirror modification H3K27me3 (Fig. 4E,F) along the whole gene. Loss of the H3K36me3 methyltransferase in *sdg8* confers early flowering and low *FLC* expression (26-28). Combination of *fca* with *sdg8* results in an *FLC* level and profile of total RNA across intron1 similar to that in Col (Fig. 5A, Fig. S8). Therefore, loss of SDG8-directed H3K36me3 is also likely to coordinately reduce Pol II initiation and elongation rates at *FLC*. Taken together, our data suggest that activities downstream of antisense processing act antagonistically to SDG8 function, leading to coordinated changes in initiation and elongation at *FLC* (Fig. 5B).

Discussion

Understanding how flowering time in plants is regulated has led into a detailed mechanistic dissection of the regulation of the *Arabidopsis thaliana* floral repressor *FLC*. Genetic screens have identified RNA processing factors that target antisense transcripts of *FLC* and histone modifiers as important components quantitatively repressing *FLC* expression. Here, using a combination of mathematical modeling and experiments, we show *FLC* regulation involves coordination of transcription initiation with elongation. This may be a general feature of gene regulation as evidenced by genome-wide correlations between gene expression,

gene body Pol II levels and Pol II elongation rates found in yeast and mammalian cells (14, 29).

How Pol II initiation and elongation are coordinated is still unclear. In *E. coli*, newly initiated RNA Polymerases can facilitate elongation of the leading Polymerase (30). Such a mechanism is unlikely to be the case at *FLC,* since *FLC* is not highly expressed even in its active state (as compared to *Actin*). Elongation is likely influenced by Pol II CTD modifications and the chromatin state (31, 32), both directly through nucleosome turnover dynamics and indirectly via differential recruitment of elongation factors. In *Arabidopsis*, elongation factor TFIIS is required for elongation of many genes but a *tfIIS* mutant does not show changed *FLC* expression (10, 33, 34). However, *FLC* expression is particularly sensitive to reduced amounts of the histone chaperone FACT (35), so it will be interesting to test if FACT is required for the fast elongation observed in *fca-9* and the coordination mechanism. We have found here that FLD recruitment, changed H3K4me2 and the resulting changes in H3K36me3 at *FLC* are likely important for this coordination. Our analysis of SDG8 suggests that H3K36me3 is essential to maintain both a fast initiation and elongation rate at *FLC* (Fig. 5B). We therefore propose that changed histone modifications actively influence *FLC* regulation and are not just a reflection of transcription.

Our results raise the question whether there is a general need to coordinate transcription initiation and elongation. Control of gene expression may necessitate such coordination as, for instance, a slow elongation rate relative to initiation would cause an accumulation of Pol II at the promoter that would limit the number of additional Pol II molecules that can initiate through occlusion (36). Such a limit might become even more stringent due to bursty initiation or Pol II pausing/backtracking during elongation (37). Furthermore, antisense transcription might induce a limit on initiation rates in order to prevent the occurrence of TI (38). However, 5' pausing of Pol II is not a feature at *FLC* (as shown by the absence of a 5' peak in Pol II ChIP), arguing against occlusion effects. The expression of sense and antisense is positively correlated at *FLC*, arguing against a major role for TI. Instead we suggest that altered elongation rates reinforce selection of different antisense isoforms, which can then recruit different chromatin regulators to the gene, thereby modulating coordinated transcription initiation and elongation (Fig. 5B). An important question now is to understand how far the lessons from *FLC* reflect regulation mechanisms both genome- and organism-wide. Coordination between initiation and elongation could generally enhance transcription efficiency, potentially to minimize transcription-associated genome instability (39). Modulation of the deposition of different histone modifiers by non-coding transcripts may be a general mechanism to coordinately affect Pol II initiation and elongation and thus quantitatively modulate transcriptional output.

Materials and Methods

Experimental procedures and mathematical modeling can be found in the Supporting Information.

Acknowledgements.

This work was supported by BBSRC grant BB/K007203/1 (MH,CD), BBSRC Institute Strategic Program GRO (BB/J004588/1) and VSBfonds Scholarship and Prins Bernhard Cultuurfonds Scholarship (RI). We thank Hervé Vaucheret for providing *xrn* seeds and Robert Sablowski for comments on the manuscript. We thank Dean and Howard group members for discussions.

1. Smolle M, Workman JL (2013) Transcription-associated histone modifications and cryptic transcription. *Biochim Biophys Acta* 1829(1):84-97.
2. Selth LA, Sigurdsson S, Svejstrup JQ (2010) Transcript Elongation by RNA Polymerase II. *Annu Rev Biochem* 79:271-293.
3. Sheldon CC, *et al.* (1999) The *FLF* MADS box gene: a repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. *Plant Cell* 11(3):445-458.
4. Michaels SD, Amasino RM (1999) *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11(5):949-956.
5. Crevillen P, Dean C (2011) Regulation of the floral repressor gene *FLC*: the complexity of transcription in a chromatin context. *Curr Opin Plant Biol* 14(1):38-44.
6. Liu F, Marquardt S, Lister C, Swiezewski S, Dean C (2010) Targeted 3' processing of antisense transcripts triggers *Arabidopsis FLC* chromatin silencing. *Science* 327(5961):94-97.
7. Hornyik C, Terzi LC, Simpson GG (2010) The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. *Dev Cell* 18(2):203-213.
8. Ietswaart R, Wu Z, Dean C (2012) Flowering time control: another window to the connection between antisense RNA and chromatin. *Trends Genet* 28(9):445-453.
9. Liu F, *et al.* (2007) The *Arabidopsis* RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate *FLC*. *Mol Cell* 28(3):398-407.
10. Marquardt S, *et al.* (2014) Functional consequences of splicing of the antisense transcript *COOLAIR* on *FLC* transcription. *Mol Cell* 54(1):156-165.
11. Wang ZW, Wu Z, Raitskin O, Sun Q, Dean C (2014) Antisense-mediated *FLC* transcriptional repression requires the P-TEFb transcription elongation factor. *Proc Natl Acad Sci U S A* 111(20):7468-7473.
12. Ehrensberger AH, Kelly GP, Svejstrup JQ (2013) Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. *Cell* 154(4):713-715.
13. Brody Y, *et al.* (2011) The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol* 9(1):e1000573.
14. Danko CG, *et al.* (2013) Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol Cell* 50(2):212-222.
15. Fuchs G, *et al.* (2014) 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biol.* 15(5):R69.
16. Singh J, Padgett RA (2009) Rates of in situ transcription and splicing in large human genes. *Nat Struct Mol Biol* 16(11):1128-1133.
17. Bentley DL (2014) Coupling mRNA processing with transcription in time and space. *Nat Rev Genet* 15(3):163-175.
18. Ameur A, *et al.* (2011) Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat Struct Mol Biol* 18(12):1435-1440.
19. Hesselberth JR (2013) Lives that introns lead after splicing. *Wiley Interdiscip Rev RNA.* 4(6):677-691.
20. Mason PB, Struhl K (2005) Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. *Mol Cell* 17(6):831-840.
21. Hazelbaker DZ, Marquardt S, Wlotzka W, Buratowski S (2013) Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. *Mol Cell* 49(1):55-66.
22. Cheng Y, Kato N, Wang W, Li J, Chen X (2003) Two RNA binding proteins, HEN4 and HUA1, act in the processing of *AGAMOUS* pre-mRNA in *Arabidopsis thaliana*. *Dev Cell* 4(1):53-66.
23. Brannan K, *et al.* (2012) mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol Cell* 46(3):311-324.
24. Jimeno-Gonzalez S, Haaning LL, Malagon F, Jensen TH (2010) The yeast 5'-3' exonuclease Rat1p functions during transcription elongation by RNA polymerase II. *Mol Cell* 37(4):580-587.
25. Gy I, *et al.* (2007) *Arabidopsis* FIERY1, XRN2, and XRN3 are endogenous RNA silencing suppressors. *Plant Cell* 19 (11):3451-3461.
26. Ko JH, *et al.* (2010) Growth habit determination by the balance of histone methylation activities in *Arabidopsis*. *EMBO J* 29(18):3208-3215.
27. Kim SY, *et al.* (2005) Establishment of the vernalization-responsive, winter-annual habit in *Arabidopsis* requires a putative histone H3 methyl transferase. *Plant Cell* 17(12):3301-3310.
28. Yang H, Howard M, Dean C (2014) Antagonistic Roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at *Arabidopsis FLC*. *Curr Biol* 24(15):1793-1797.
29. Mayer A, *et al.* (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* 17(10):1272-1278.
30. Epshtein V, Nudler E (2003) Cooperation between RNA polymerase molecules in transcription elongation. *Science* 300(5620):801-805.
31. Jonkers I, Kwak H, Lis JT (2014) Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* 3: e02407.
32. Weber CM, Ramachandran S, Henikoff S (2014) Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell* 53(5):819-830.
33. Van Lijsebettens M, Grasser KD (2014) Transcript elongation factors: shaping transcriptomes after transcript initiation. *Trends Plant Sci* 19(11):717-726.
34. Dolata J, *et al.* (2015) NTR1 is required for transcription elongation checkpoints at alternative exons in *Arabidopsis*. *EMBO J* 34(4):544-558.
35. Lolas IB, *et al.* (2010) The transcript elongation factor FACT affects *Arabidopsis* vegetative and reproductive development and genetically interacts with HUB1/2. *Plant J* 61(4):686-697.
36. Core LJ, *et al.* (2012) Defining the status of RNA polymerase at promoters. *Cell Rep* 2(4):1025-1035.
37. Churchman LS, Weissman JS (2011) Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* 469 (7330):368-373.
38. Hobson DJ, Wei W, Steinmetz LM, Svejstrup JQ (2012) RNA polymerase II collision interrupts convergent transcription. *Mol Cell* 48(3):365-374.
39. Saponaro M, *et al.* (2014) RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. *Cell* 157(5):1037-1049.
40. Wuarin J, Schibler U (1994) Physical isolation of nascent RNA chains transcribed by RNA polymerase II: evidence for cotranscriptional splicing. *Mol Cell Biol* 14(11):7219-7225.
41. Almada AE, Wu XB, Kriz AJ, Burge CB, Sharp PA (2013) Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* 499(7458):360-363.
42. Tippmann SC, *et al.* (2012) Chromatin measurements reveal contributions of synthesis and decay to steady-state mRNA levels. *Mol Syst Biol* 8: 593.
43. Gray JM, *et al.* (2014) SnapShot-Seq: A method for extracting genome-wide, in-vivo mRNA dynamics from a single total RNA sample. *Plos One* 9(2).