Transcriptional read-through of the long non-coding RNA SVALKA governs plant cold acclimation

Peter Kindgren, Ryan Ard, Maxim Ivanov & Sebastian Marquardt

Most DNA in the genomes of higher organisms does not encode proteins, yet much is transcribed by RNA polymerase II (RNAPII) into long non-coding RNAs (lncRNAs). The biological significance of most lncRNAs is largely unclear. Here, we identify a lncRNA (SVALKA) in a cold-sensitive region of the Arabidopsis genome. Mutations in SVALKA affect CBF1 expression and freezing tolerance. RNAPII read-through transcription of SVALKA results in a cryptic lncRNA overlapping CBF1 on the antisense strand, termed asCBF1. Our molecular dissection reveals that CBF1 is suppressed by RNAPII collision stemming from the SVALKA-asCBF1 lncRNA cascade. The SVALKA-asCBF1 cascade provides a mechanism to tightly control CBF1 expression and timing that could be exploited to maximize freezing tolerance with mitigated fitness costs. Our results provide a compelling example of local gene regulation by lncRNA transcription having a profound impact on the ability of plants to appropriately acclimate to challenging environmental conditions.
RNA Polymerase II (RNAPII) transcription in genomes results in the production of many long non-coding RNAs (lncRNAs)\(^1\). The functional significance of resulting lncRNA molecules is actively debated even though biological roles have been identified for an increasing number of examples\(^2\text{-}^3\). Expression of lncRNAs is remarkably specific to the environmental condition, tissue or cell type, arguing for roles of lncRNAs in regulation\(^4\text{-}^7\). In addition to functions carried out by IncRNA molecules, the process of transcribing non-coding DNA sequences can by itself be regulatory in many systems\(^8\text{-}^9\). Non-coding DNA regions in genomes can therefore affect gene expression by different mechanisms that need to be resolved experimentally.

Sessile organisms respond to changing environmental conditions by the regulation of gene expression. Early events in the Arabidopsis cold response include rapid transcriptional up-regulation of the intron-less C-repeat/dehydration-responsive DNA regions in genomes can therefore affect gene expression by different mechanisms that need to be resolved experimentally. The plant response following long exposure to cold temperatures in plants are mediated by lncRNAs. However, it is currently unresolved if early responses to cold transcriptional read-through of the lncRNA \(\text{fi}^{4.23}\). This cold-triggered cascade of lncRNAs establishes long-lasting and stable repression of the \(\text{FLC}\) gene and serves as a paradigm for gene repression by lncRNA\(^23\). However, it is currently unresolved if early responses to cold temperatures in plants are mediated by lncRNAs.

Here, we map transcriptional start sites in Arabidopsis during early responses to cold temperatures. We identify a cascade of two lncRNAs that fine-tune the expression of the \(\text{CBF1}\) gene. Transcriptional read-through of the lncRNA \(\text{SVALKA}\) results in the expression of a cryptic antisense \(\text{CBF1}\) lncRNA (\(\text{asCBF1}\)). \(\text{asCBF1}\) transcription results in RNAPII collision to limit the expression of full-length \(\text{CBF1}\). This work extends the biological roles of lncRNAs during cold acclimation and it provides an elegant mechanism that achieves rapid dynamic regulation of environmentally sensitive gene expression.

Results

Identification of the lncRNA \(\text{SVALKA}\). To identify transcription initiation events that respond to cold temperature in Arabidopsis, we performed transcription start site (TSS)-sequencing (TSS-seq)\(^10\text{-}^11\). We obtained 184 million raw reads from two biological replicates at 22°C and two biological replicates at 3 h of 4°C. From these, we called 14250 TSS clusters. Our analyses revealed 489 down-regulated and 1404 up-regulated TSSs in response to 3 h at 4°C (Fig. 1a, Supplementary Data 1, Supplementary Data 2). Most differentially used TSSs were located in promoter regions corresponding to changed promoter usage in fluctuating environments. Of the up-regulated genes (\(\pm 1200 \text{bp of the major TSS}\), the \(\text{CBF}\) genes were upregulated 100–400 fold making the \(\text{CBF}\) genomic region by far the most cold-responsive region in the genome (Fig. 1b). Interestingly, we detected a cold-responsive lncRNA, transcribed on the antisense strand between \(\text{CBF3}\) and \(\text{CBF1}\) (Fig. 1c) that we named \(\text{SVALKA}\) (SVK, for short). We found two TSSs by Rapid Amplification of cDNA Ends (RACE) that corresponded to the peaks observed with TSS-seq (Fig. 1d). Our 3’ RACE analysis also supported the existence of alternative SVK isoforms. Notably, we detected no transcript overlapping the downstream \(\text{CBF1}-3’\text{-UTR}\) (Supplementary Figure 1, Supplementary Figure 2). RT-qPCR revealed that more \(\text{SVK}\) is generated from the \(\text{CBF1}\)-proximal promoter than from the distal promoter (Supplementary Figure 3), consistent with TSS-seq analyses. A time series of cold exposure (4°C) with samples taken every 4 h revealed the expression pattern of \(\text{SVK}\) and \(\text{CBF1}\) (Fig. 1e). \(\text{CBF1}\) expression peaked at 4 h followed by a decrease that reached a relatively stable level after 12–16 h, a pattern in congruence with earlier studies\(^26\). In contrast, SVK showed gradual increase in expression from 4 h to reach a stable maximum level after 12–16 h.

\(\text{SVK}\) represses \(\text{CBF1}\) and promotes cold acclimation. The anti-correlation between \(\text{CBF1}\) and \(\text{SVK}\) and their genomic proximity suggested that \(\text{SVK}\) could be involved in \(\text{CBF1}\) repression. To test this, we constructed two distinct \(\text{LUCIFERASE (LUC)}\) reporter lines of \(\text{CBF1}\) with different termination sequences. We expected any repressive role of \(\text{SVK}\) to be present in the lines with endogenous 3’-sequences (\(\text{CBF1}\)-\(\text{SVK}\)). In contrast, we expected effects of \(\text{SVK}\) to be absent in the lines with the terminator sequences of the \(\text{Nopaline Synthase (NOS)}\) gene that shows low antisense transcription\(^4\) (\(\text{CBF1}\)-\(\text{TNOS}\)). We subjected three independent lines from each construct to 4°C for 0–12 h and subsequently measured LUC activity (Fig. 2a, b). The three \(\text{CBF1}\)-\(\text{TNOS}\) lines showed increased LUC activity compared to the \(\text{CBF1}\)-\(\text{SVK}\) lines. These data demonstrated that \(\text{SVK}\) represses LUC activity. We wanted to confirm these findings by measuring antisense transcription from the 3’-end of the \(\text{LUC}\) constructs. Therefore, we designed a qPCR based strategy to measure antisense transcription that may be initiated from the endogenous \(\text{CBF1}-3’\text{-UTR}\) or the \(\text{TNOS}\) terminator\(^4\). We detected antisense transcription from the \(\text{TNOS}\) sequence in two independent \(\text{CBF1}\)-\(\text{TNOS}\) lines (Supplementary Figure 4). However, antisense transcription from \(\text{TNOS}\) was not cold-responsive, arguing against a cold-induced effect of \(\text{LUC}\) expression. In contrast, we detected cold-induced antisense akin to \(\text{SVALKA}\) transcription from the endogenous \(\text{CBF1}-3’\text{-UTR}\) in the \(\text{CBF1}\)-\(\text{SVK}\) lines (Supplementary Figure 4). These data suggest that cold-induced antisense transcription could be involved in regulating endogenous \(\text{CBF1}\) expression.

We next characterized Arabidopsis Transfer DNA (T-DNA) lines to further dissect the function of the \(\text{CBF1}\) 3’-region. T-DNA lines insert large sequences of DNA (often 4–5 kb) in the Arabidopsis genome that disrupt the genomic context at the insertion site\(^27\). We isolated three informative T-DNA insertion mutants: (1) a line that disrupted \(\text{SVK}\) (\text{svk-1}), (2) a line that increased the distance of \(\text{SVK}\) transcription from \(\text{CBF1}\) (\text{uncoupling svalka-1, uns-1}), and (3) a line that showed \(\text{SVK}\) over-expression (\(\text{SVK OE}\)) (Fig. 2c, Supplementary Figure 2, 5). The \(\text{SVK}\) over-expression originates from strong promoters within the T-DNA that continued over the insertion border (Supplementary Figure 5a-b). We could not detect any other T-DNA driven transcription in the \text{svk-1} or \text{uns-1} mutants (Supplementary Figure 5c-d). We quantified \(\text{SVK}\) and \(\text{CBF1}\) expression during cold acclimation11. The negative gene is characterized by a transient peak of maximal transcription during cold exposure1. The negative activation pattern of the lncRNA \(\text{fi}^{4.23}\). For example, over-expression of \(\text{CBF1}\) results in increased tolerance to freezing temperatures, but also severely reduces biomass\(^20\). Therefore, the expression of the endogenous \(\text{CBF1}\) gene is characterized by a transient peak of maximal expression during cold acclimation\(^1\). The negative fitness costs of both increased and decreased \(\text{CBF}\) expression highlight the importance of their tight regulation, yet the specific mechanisms that tightly control \(\text{CBF}\) expression during the early response to cold are not fully understood.

The plant response following long exposure to cold temperatures is associated with numerous lncRNAs that associate with the \(\text{FLOWERING LOCUS C (FLC)}\) gene. The antisense lncRNA \(\text{COOLAIR}\) is induced after 2 weeks of cold, and initiates chromatin repression of the \(\text{CBF1}\) gene. This work extends the biological roles of lncRNAs during cold acclimation and it provides an elegant mechanism that achieves rapid dynamic regulation of environmentally sensitive gene expression.
expression by RT-qPCR. svk-1 showed reduced expression of SVK while the uns-1 mutants showed slightly elevated levels of SVK compared to wild-type. As expected, the SVK OE mutant showed greatly increased level of SVK (Fig. 2d). Interestingly, we detected CBF1 mis-regulation in all three mutants (Fig. 2e). The combination of high level of SVK expression and increased CBF1 expression in uns-1 compared to wild-type argues against a trans-acting function for the SVK lncRNA (Fig. 2d, e). In this mutant, SVK is expressed 4–5 kb away (compared to 110 bp for wild-type) from CBF1 but CBF1 expression is still increased compared to wild-type. Both the uns-1 and svk-1 mutants showed an increase of CBF1 expression in response to cold exposure, while the SVK OE mutant lead to decreased CBF1 levels, supporting the notion that SVK plays a repressive role on CBF1 mRNA levels. We could not see any effect on the neighboring CBF3 and CBF2 genes (Supplementary Figure 6). Increased CBF1 levels in the uns-1 and the svk-1 mutants lead to greater induction of CBF-activated COR genes (Fig. 3a). These findings suggested that SVK transcription could have a biologically important role and affect cold acclimation and freezing tolerance. We conducted an electrolyte leakage test of wild-type, svk-1 and uns-1 plants to test the role of SVK in freezing tolerance (Fig. 3b). In this test, leaf discs of non-acclimated and cold-acclimated plants are in contact with water and exposed to decreasing freezing temperatures and measured for leakage of electrolytes (i.e., plasma membrane disruption). Surviving cells keep their electrolytes within the cell. Therefore, measuring electrolyte leakage is a measurement of how well the cells can survive freezing temperatures. Cold-acclimated svk-1 and uns-1 showed a higher freezing tolerance compared to wild-type. Wild-type plants showed an electrolyte leakage of 50% at −6.5 °C (±0.2 °C, 95% CI) compared to svk-1 (−7.7 °C, ±0.2 °C) and uns-1 (−7.9 °C, ±0.3 °C). The increase in freezing tolerance is remarkable and is similar to that of the myb15 mutant, a transcription factor that directly suppresses CBF expression. Together, our findings suggest that the genomic region encoding SVK represses cold-induced CBF1 expression and has a biologically relevant effect on cold acclimation and freezing tolerance in Arabidopsis.
Read-through transcription of SVK results in asCBF1. To better understand the mechanism of SVK repression, we revisited the effects of the uns-1 insertion. uns-1 is inserted downstream of SVK yet results in equivalent cold-related defects compared to svk-1. A possible explanation for the effects seen in the svk-1 and uns-1 mutants would be that read-through transcription of CBF1 is reduced by the T-DNA insertions that may increase the stability of the CBF1 mRNA. The Arabidopsis 5′–3′ exonuclease XRN3 is associated with the removal of uncapped nuclear transcripts that are expected for read-through transcripts derived from RNAPII termination. Consistently, we find increased CBF1 read-through transcription in the xrn3–3 mutant (Supplementary Figure 7). However, we could not detect decreased CBF1 read-through in the svk-1 and uns-1 mutants, arguing against increased mRNA stability linked to increased transcriptional termination efficiency in these mutants (Supplementary Figure 7). Another possibility that could reconcile our findings would be if SVK was promoting transcription of a cryptic antisense transcript into the CBF1 gene body, since such a transcript would be disrupted in uns-1 and svk-1. To test this hypothesis, we carefully examined the presence of an antisense transcript mapping to the 3′-UTR of CBF1 (Fig. 4a, b). HEN2 is part of the nucleoplasmic 3′ to 5′ exosome responsible for degrading many types of non-coding RNA transcripts. We used a high resolution time series of cold exposed samples of wild type and hen2-2 to enable the detection of cryptic antisense transcripts. A CBF1 antisense transcript (asCBF1) corresponding to roughly 250 nt was detectable in the hen2-2 mutant, yet not in the wild type (Fig. 4a, b).
acclimated (left panel) and cold-acclimated (right panel) WT (black), contact with deionized water in a test tube and exposed to replicates. Source data are provided as a Source Data nitorgen/total electrolyte content) after freezing test compared to the total electrolyte content (±standard deviation) from at least three biological allowed to thaw and measured again for electrolytes. Each marker represents the mean electrolyte leakage (electrolyte content before exposure to liquid at the indicated temperatures and the solution in the tubes was measured for electrolytes. Subsequently, each tube was submersed in liquid nitrogen determined with Student’s t-test (*p < 0.05, **p < 0.01, ***p < 0.001). Source data are provided as a Source Data file. b Electrolyte leakage in non-acclimated (left panel) and cold-acclimated (right panel) WT (black), svk-1 (light grey) and uns-1 (dark grey) plants. Leaf discs of each genotype were in contact with deionized water in a test tube and exposed to −2 °C for 1 h followed by a decreased temperature at a rate of 2 °C/h. Samples were taken out at the indicated temperatures and the solution in the tubes was measured for electrolytes. Subsequently, each tube was submersed in liquid nitrogen allowed to thaw and measured again for electrolytes. Each marker represents the mean electrolyte leakage (electrolyte content before exposure to liquid nitrogen/total electrolyte content) after freezing test compared to the total electrolyte content (+standard deviation) from at least three biological replicates. Source data are provided as a Source Data file

Supplementary Figure 8a), asCBF1 was also detected in additional mutants disrupting the nuclear exosome (Supplementary Figure 8b). However, we could not detect asCBF1 in the sop1-5 ( suppressor of pas2) mutant that have an increase of a subset of HEN2-dependent degradation targets or the trl-1 mutant, which is disrupted for TRL1 (TRF4/5-like), a factor involved in 3′-processing of tRNA in the nucleolus (Supplementary Figure 8b). Together, these results suggest asCBF1 transcript is a nucleoplasmic exosome target with expression levels correlated to SVK. The size of asCBF1 suggests it is either a SVK read-through transcript that is cleaved at the poly-(A) signal or a new transcription initiation event. To confirm that asCBF1 depends on SVK transcription, we crossed the svk-1 and uns-1 mutants to hen2-2 (Fig. 4c). In both double mutants, the asCBF1 RNA was lost, confirming that asCBF1 transcription depends on SVK.

Transcription in higher eukaryotes occurs beyond the poly-(A) termination signal before nascent transcript cleavage is triggered. Nascent transcript cleavage creates the 3′-end of the full-length mRNA substrate for poly-adenylation, and a free 5′-end is attacked by 5′ to 3′ exonucleases eliciting transcription termination according to the torpedo model. To detect any read-through transcription from SVK, we purified RNA associating with RNAPII complexes (Supplementary Figure 9a). To isolate nascent RNA, we used a stable Arabidopsis line with a FLAG-tagged version of NRPB2 (i.e., the second largest subunit of RNAPII). After IP, NRPB2 co-eluted with NRPB1, confirming that we purified intact RNAPII complexes (Supplementary Figure 9b-d). Subsequently, we isolated the RNA attached to the purified RNAPII complexes. We profiled this RNA for the presence of SVK transcripts extending into CBF1 as antisense transcripts with qPCR (Fig. 4d). We detected quantifiable levels of a read-through product after 4 h of cold exposure and increased expression after 8 h in the nascent RNA fraction. Consistent with asCBF1 being a product of read-through following SVK 3′-end formation and cleavage, we could not detect the SVK-asCBF1 read-through transcript in total RNA (i.e. steady state) or in a Mock-IP control samples (i.e. IP with a non-FLAG tagged line). Detection in nuclear exosome mutants and nascent transcript preparation gives asCBF1 characteristics of a cryptic lncRNA that is quickly cleaved and degraded in wild-type. XRN3 is thought to represent the 5′-to-3′ torpedo exonuclease mediating transcriptional termination in Arabidopsis. However, we did not detect
the asCBF1 RNA or CBF1 mRNA mis-regulation in the xrn3-3 mutant (Supplementary Figure 10). These data suggest that transcriptional termination of SVK read-through is either XRN3 independent, or the hypomorphic xrn3-3 allele does not disrupt XRN3 function strong enough to detect SVK termination defects. In summary, our findings revealed asCBF1 as a HEN2-dependent

Fig. 4 SVALKA transcription mediates transcription activity antisense to CBF1. a Representative Northern blot of a cold exposure time series of Col-0 (WT) and hen2-2. The probe used for asCBF1 is shown in b. Blots were repeated with three biological replicates with similar results. UBI is used as loading control. Uncropped blots can be found in the Source Data file. b Graphical representation of the CBF1-SVK genomic region. The probe for asCBF1 is shown in red. RT was done with an oligo-linked primer to ensure strand-specificity and generated cDNA was used in c. c RT-qPCR of asCBF1 after cold exposure in hen2-2 and the double mutants, hen2-2uns-1, and hen2-2svk-1. The RT-primer and the qPCR primers are shown in the graphical representation above the graph. Bars represent mean (white: 4 h 4 °C, black: 8 h 4 °C, ±SEM) from three biological replicates (rings). The relative level of asCBF1 was normalized to the level in hen2-2 after 4 h of 4 °C. Statistically significant differences were determined with Student’s t-test (**p < 0.001). Source data are provided as a Source Data file. d) RT-qPCR of SVK read-through transcripts. Hundred nanograms of RNA from an RNAPII-IP (i.e. RNA attached to RNAPII) or total RNA was converted to cDNA with an oligo-linked primer. The RT-primer and qPCR primers used can be seen in the graphical representation above the graph. The RT-primer annealed to a sequence downstream of the PolyA signal of SVK. A read-through transcript can be detected in the nascent RNA samples with an increased expression after 8 h of cold exposure. No signal could be seen in the Mock-IP or total RNA sample. Each bar represents the mean (white: 4 h 4 °C, black: 8 h 4 °C, ±SEM) from three biological replicates (rings). The relative level of read-through transcripts was normalized to the level in the nascent sample after 4 h of 4 °C. Statistically significant differences were determined with Student’s t-test (**p < 0.001). Source data are provided as a Source Data file.
“cryptic” antisense transcript that depends on SVK read-through transcription during cold to repress CBF1 induction.

**CBF1 expression is limited by RNAPII collision.** Two molecular mechanisms could explain the observed SVK-mediated effects on CBF1 expression: (1) transcription of asCBF1 activates the siRNA pathway to trigger repression, or (2) the act of asCBF1 transcription itself causes sense/antisense competition and RNAPII collision\(^{35}\). To test if siRNA were involved, we performed cold exposure of several different mutants in various siRNA pathways and probed for CBF1 mRNA levels (Supplementary Figure 11a). All the mutants failed to show increased CBF1 levels (Supplementary Figure 11b), ruling out the siRNA model. A RNAPII collision model predicts a discrepancy of transcription between the 5′-end and the 3′-end of CBF1 caused by stalled RNAPII transcription complexes in the 3′-end of CBF1\(^{36}\). Our protocol to isolate nascent RNA enabled us to detect strand-specific sites of RNAPII accumulation (i.e., potential pause sites) with RT-qPCR (Fig. 5b, Supplementary Figure 9). The 5′-end probe (Probe 1 in Fig. 5c) showed no significant difference, suggesting that RNAPII occupancy early in CBF1 elongation was indistinguishable between 4 and 8 h of 4°C. The probe in the 3′-end of the CBF1 exon (Probe 2 in Fig. 5c) showed that RNAPII occupancy was significantly increased over this region of CBF1 after 8 h compared to 4 h. Intriguingly, the increased RNAPII occupancy at the 3′-end of the CBF1 exon was not detected in the 3′-UTR (Probe 3 in Fig. 5c). These data are consistent with the collision model that predicts RNAPII complexes are pre-maturely terminated before they have time to fully transcribe a full-length CBF1 transcript.

The 3′-UTR probe assays the steady state level of full-length CBF1 mRNA (Fig. 2e). The steady state CBF1 RNA levels after 4 and 8 h of cold exposure showed lower levels of CBF1 transcripts containing the middle (probe 2 in Fig. 5c) and the end (probe 3 in Fig. 5c) than the beginning of CBF1 (probe 1 in Fig. 5c). These data confirm that the major decrease after 8 h of cold exposure in full-length CBF1 mRNA levels relate to differences in the 3′-end. In agreement with these results, we could detect high amounts of pre-maturely terminated CBF1 when sense transcripts are probed in the 5′-end compared to the 3′-end with Northern blot (Supplementary Figure 12). This effect was exaggerated in the hen2-2 mutant, suggesting that many pre-terminated CBF1 events are rapidly degraded by the nuclear exosome. Probes for antisense transcripts of SVK showed an increasing RNAPII occupancy over the SVK transcription unit following longer exposure to cold (Probe 6 in Fig. 5d). We detected SVK read-through transcripts in the nascent fraction (Probe 5 in Fig. 5d). No products were detected with the probe further downstream (Probe 4 in Fig. 5d). These data support that RNAPII complexes are terminated before transcription reaches further into the CBF1 exon. This is also consistent with the size of the asCBF1 transcript (Fig. 4a). We complemented these results with total RNAPII quantitative Chromatin Immuno-Precipitation (qChIP) analyses following exposure to 4°C. Reminiscent of the results from analyzing nascent transcripts, we detected a higher RNAPII occupancy over the CBF1 exon after 8 h compared to 4 h of 4°C. In contrast, the probe at the promoter of CBF1 (pCBF1 in Fig. 5a–e) showed a decrease of RNAPII occupancy after 8 h. These results strongly suggest that RNAPII complexes are stalled in the 3′-end of CBF1 since full-length CBF1 mRNA levels decrease after 8 h compared to 4 h of cold exposure (Figs 1e, 5c). qChIP probes for SVK showed higher use of the SVK promoter following cold exposure that was also seen for the SVK transcription unit (Fig. 5f). Another prediction of the collision model is that increased RNAPII occupancy over the CBF1 exon in wild-type should be abolished in the uns-1 mutant, since no collision would occur due to the absence of asCBF1 transcription. Our hypothesis was tested and confirmed by qChIP for total RNAPII over the CBF1 exon in wild type and uns-1 since no increase of RNAPII occupancy was observed in the uns-1 mutant (Fig. 5g).

Collectively, these results imply the presence of stalled RNAPII complexes over the 3′-end of CBF1 that were correlated with increased asCBF1 expression. In conclusion, our results showed that read-through of a lncRNA triggers head-to-head RNAPII collision over the CBF1 gene body leading to regulated termination of CBF1 transcription and a decrease of full-length CBF1 mRNA levels (Fig. 5h). This study illustrates how lncRNA read-through transcription limits peak expression of gene expression to promote an appropriate response to environmental change.

**Discussion**

Environmental change alters RNAPII transcription in genomes so that organisms can adequately respond to new conditions\(^{3,39}\). These transcriptional responses include high RNAPII activity in non-coding regions, resulting in lncRNA transcription. On the one hand, lncRNA molecules may provide function, for example through riboswitch regulation\(^{40,41}\). On the other hand, the process of lncRNA transcription by RNAPII can affect the regulation of gene expression in the vicinity of target genes\(^{42}\). A curious result of widespread lncRNA transcription is that genes are often transcribed on both DNA strands, resulting in mRNAs from the sense strand, and lncRNA from the antisense strand\(^{43}\).

Antisense transcription is detected in many genomes, and genome-wide analyses suggest that these phenomena can be either positively or negatively correlated with the corresponding sense transcript\(^{44}\). The positive correlations between sense and antisense expression represent a puzzling conundrum, as this relationship poses the question of how two converging RNAPII enzymes transcribing the same DNA can move past each other\(^{38}\). Sense and antisense transcription pairs on cell population level have been resolved as binary expression states for the individual transcripts at single-cell level\(^{45}\). Perhaps alternative expression states of sense or antisense lncRNA transcription in individual cells offers an explanation for why examples of gene repression by RNAPII collision are relatively rare. Overall, antisense transcription appears to be a common strategy to affect gene repression\(^{44}\).

Our study uncovers how the interplay between two lncRNAs limits the production of maximal mRNA levels of the *Arabidopsis CBF1* gene by an RNAPII collision mechanism in the 3′-end of CBF1. Cold-induced transcription initiation in the CBF locus triggers expression of a cascade of antisense transcripts that fine-tunes the level of CBF1 mRNA, thereby representing a negative feedback required for plants to appropriately acclimate to low temperatures.

The identification of interplay between a stable lncRNA SVALKA and a cryptic lncRNA asCBF1 acting together in gene repression represents an important aspect of our study. The cryptic lncRNA asCBF1 is detectable after 6 h of cold treatment in nuclear exosome mutant backgrounds. asCBF1 is likely representative of additional lncRNAs with regulatory roles that are challenging to study owing to their cryptic and environmental specific expression characteristics. Our findings of multiple lncRNAs acting together is in line with recent findings suggesting a coordinated effect of multiple lncRNAs to regulate meiosis in yeasts\(^{46,47}\). The combinations of multiple lncRNAs amplify the interfaces with regulatory potential and may increase precision of target regulation. Since asCBF1 results from RNAPII read-through transcription of SVK, our findings highlight the process of transcriptional termination of lncRNAs for regulation.
Selective transcriptional termination of the *Arabidopsis* lncRNA COOLAIR determines *FLC* expression states by a chromatin-based mechanism\(^2\). A similar mechanism is more difficult to elucidate in this case due to the relatively small size of the *CBF1* gene body (i.e. 946 bp). Moreover, the class of budding yeast stable uncharacterized transcripts (SUTs) with environment-
specific expression is characterized by inefficient termination resulting in read-through transcription into neighboring genes. Consistently, regulation of yeast gene expression by read-through transcription of SUTs is well-documented. Extended SUT species link transcripts to create interdependence of loci to achieve coordinated expression changes. The effects of the SVK-asCBFI1 circuitry support the idea that cascades of IncRNA transcription provide a mechanism to coordinate local regulation. IncRNA cascades increase the spacing distance between initial IncRNA transcription events and the target gene, in contrast to a straightforward antisense IncRNA. Multiple cascading transcripts provide more room for independent regulatory inputs to achieve precision in regulation. The topology of the Arabidopsis genome is thought to be gene-centric with local gene loops rather than larger topologically associated domains (TADs). We would expect extensive co-regulation if sense and antisense IncRNA transcription were part of the same gene loop. Indeed, alternative gene loop formation is triggered by transcription of the IncRNA APOLO to regulate expression of the PINOID gene. Read-through transcription of the IncRNA SVALKA may provide a mechanism to invade and break positive feedback mechanisms reinforcing CBF1 transcription. As IncRNA transcription in yeast and mammalian cells correlate with the boundaries of chromatin domains, roles of IncRNA in communicating regulation across domain boundaries may be common despite differences in genomic topologies.

The entire transcripts in the CBF cluster increase expression in response to short-term cold temperatures. The discovery of the SVK-asCBFI1 IncRNA cascade illustrates how transcriptional activation of this region is associated with an inbuilt negative feedback mechanism that limits the timing of maximal CBF1 expression by triggering RNAPII collision. While CBF genes promote freezing tolerance, excess CBF gene expression is associated with fitness penalties. In this regard, the SVK-asCBFI1 cascade provides a mechanism to tightly control maximal CBF1 expression and timing that could be exploited to engineer freezing tolerance with mitigated fitness costs.

The conservation of CBF gene clusters and their role in promoting plant freezing tolerance potentially suggests that IncRNA cascades equivalent to SVK-asCBFI1 could exist in other species. The IncRNA-mediated effects of long-term cold exposure on plant gene expression have been well characterized for the FLC gene. Our research demonstrates that IncRNAs also mediate the response to short-term and transient cold exposure by timing maximal gene expression. Future research will be needed to investigate if biophysical properties of IncRNA transcription may provide a particularly effective sensor for temperature. Perhaps more likely, IncRNA functions in temperature responses will coincide with the intensity of research efforts addressing this question in plants to prepare for changing climates. In sum, our findings provide a compelling example of functional IncRNA transcription in cells to limit gene induction by triggering RNAPII collisions.

Cloning and LUC assay. Primers for cloning can be found in Supplementary Data 4. For the CBF1-Tnos construct, a fragment encompassing the CBF1 promoter and 5′UTR was amplified from genomic DNA with Phusion polynzyme (Thermo Fisher Scientific, USA) and ligated to the pENTR/D-TOPO vector (Invitrogen, USA). After sequencing to eliminate any cloning errors the pENTR/CBF1 vector was incubated with pGWBS35 vector in a LR reaction. The final vector, pGWBS35/CBF1-Tnos was transformed into Col-0 plants using floral dip method. The T1 vector was amplified from genomic DNA and fused to a CBF1-LUC fragment amplified from pGWBS35/ CBF1-Tnos. The fused fragment was ligated to pENTR/D-TOPO and sequenced. A LR reaction between pENTR/CBF1/SVK and pGWBS01 the final vector, pGWBS01/CBF1/SVK, was transformed into Col-0. Transformed seedlings were selected on Hygromycin plates and their progeny (T2 generation) was used for the LUC assay. For detection of LUC, 10 day old seedlings were treated with cold temperature (4 °C). At indicated times 5 µM D-Luciferin (ZellBio, Germany) was sprayed onto seedlings followed by dark incubation for 30–60 min. LUC was detected using a CCD camera and pixel intensity was determined in ImageJ. The average pixel intensity from at least five seedlings was used to get a mean value and used for statistical analysis (see Source Data file).

RNA extraction, northern analysis, RACE, and RT-qPCR. Total RNA was extracted using RNeasy Plant Mini Kit (Qiagen, Germany) according to manufacturer’s instructions. Northern analysis was performed as described with minor modifications. Total RNA (20 µg) was separated on a 1.2% agarose gel and transferred to a nylon membrane with UV radiation. Radiograms were detected using a phosphorimager screen (GE Healthcare, UK). RACE experiments were done with a SMARTer RACE 5′/3′ Kit (Takara, Japan) according to manufacturer’s instructions. For RT-qPCR, total RNA was DNase treated with TURBO DNase (Thermo Fisher Scientific, USA) and 1 µg DNase treated RNA was subsequently turned into cDNA as per manufacturer’s instructions with SuperScript III (Invitrogen, USA) using random primers, oligo dT or gene specific primers depending on the experiment. Diluted cDNA (1:10) was used in a PCR reaction with GoTaq qPCR Master mix (Promega, USA) and run on a CFX384 Touch instrument (Bio-Rad, USA). Data were processed in CFX manager and exported to Excel (Microsoft, USA) for analysis. Relative expression was calculated using the ΔΔCT method and at least two independent internal reference genes. All primers used in this study can be found in Supplementary Data 4. All uncropped membranes and all source data for graphs can be found in the Source Data file.

Electrolyte leakage test. Electrolyte leakage test was performed with leaf discs of non-acclimated or cold-acclimated plants in glass tubes with 200 µl of deionized water. The tubes were then transferred to a programmable bath at −2 °C (FP51, Julabo, Germany). After 1 h, the temperature was slowly decreased (−2 °C/h). Samples were taken out of the bath at designated temperatures and cooled on ice for an hour followed by 4 °C. When all samples were collected, 1.3 ml of deionized water was added and the tubes were shaken overnight at 4 °C. Electrolyte leakage was measured using a conductivity cell (CDM210, Radiometer, Denmark) and total ion content was determined in liquid nitrogen, thawed, shaken again overnight and measured for conductivity. Electrolyte leakage was determined by comparing the measured conductivity before and after the liquid nitrogen treatment. Data was fitted to a sigmoidal dose-response with GraphPad Prism and can be found in the Source Data file.

TSS-seq and bioinformatic analysis. The genome-wide distribution of TSSs in wild-type was recently mapped in Arabidopsis using 5′ CAP-sequencing. Here, we extended the read analysis to 2-weeks of cold stress (8 h at −4 °C) and micrograms of DNase-treated total RNA were treated with CIP (NEB) to remove non-capped RNA species. 5′-capped were removed using Cap-Clip (CellScript) to permit ligation of single-stranded rP5 RND adapter to 5′-ends with T4 RNA ligase 1 (NEB). Poly(A)-enriched RNAs were captured with oligo(dT) Dynabeads (Thermo Fisher Scientific) according to manufacturer’s instructions and fragmented in fragmentation buffer (50 mM Tris acetate pH 8.1, 100 mM KOAc, 30 mM MgO/A) for 5 mins at 80 °C. First-strand cDNA was generated using SuperScript III (Invitrogen) and random primers following manufacturer’s instructions. Second-strand cDNA was generated with the BioNotII F-P5 PET oligo and using Phusion High-Fidelity Polymerase (NEB) as per manufacturer’s instructions. Bioanalyzer PCR products were captured by streptavidin-coupled Dynabeads (Thermo Fisher Scientific), end repaired with End Repair Enzyme mix (NEB), A-tailed with Klenow fragment exo- (NEB), and ligated to barcoded Illumina compatible adapter using T4 DNA ligase (NEB). Libraries were amplified by PCR, size selected using Ampure XP beads (Beckman Coulter), pooled following quantification by Bioanalyzer, and sequenced in single end mode on the following flowcell: NextSeq 500/550 High Output Kit v2 (75 cycles) (Illumina). For the bioinformatics, all supplementary code for the data analysis pipelines described
below is available at [https://github.com/Maxim-Ivanov/Kindgren_et_al_2018]. The
NGS data manipulations were detailed in the 01-Processing_Scap_SEQ_data.sh pipeline.
In the custom adapter sequences (ATC/TGAT/CGCG) were trimmed from 3' ends of the Scap-Seq (TSS-Seq) reads using Trim Galore v0.4.3. Then 8 nt random barcodes (unique molecular identifiers, or UMI's) were trimmed from 5' ends of reads and appended to read names using a custom script (UMI_to_Fastq_header.py). The resultant Fastq files were aligned to TAIR10 genome using STAR v2.5.2b in the end-to-end mode. SAM files were sorted and converted to BAM using Samtools v1.2.5. Reads overlapping the rRNA, tRNA, snRNA and snorRNA genes (obtained from the Araport11 annotation) were filtered out using Bedtools v2.25.0. In addition, multimapper reads with MAPQ score below 10 were removed by Samtools. Moreover, we filtered out both PCR duplicates by analyzing the reads sharing the same 5' genomic position and removing reads with redundant UMIs (this was done using a custom script Deduplicate_BAM_files_on_UMI.py). Finally, stranded Bedgraph files were generated using Bedtools. For visualization in genomic environment, the forward and reverse Bedgraph files corresponding to the same sample were combined together and normalized to 1 million tags. For the downstream analysis in R environment, the stranded Bedgraph files were filtered for coverage ≥2x and ensured to contain only genomic intervals with 1 bp width (Expand_bedGraph_to_single_base_resolution.py). The detection of 5' tag clusters (TCS) was described in detail in the 02-Calling_Scap_SEQ_TSSR pipeline. It makes an extensive use of the CAGEEightR package [https://bioconductor.org/packages/releases/bioc/html/CAGEEightR.html]. Adjacent TSS separated by not more than 20 bp were merged into TSS clusters. The TSS clusters were annotated by intersection with various genomic features which were extracted from the TxDb.Athaliana.Rnor3.35 gene annotations. For immunoprecipitations, Protein A magnetic beads (GenScript) were used. For visualization in genomic environment, the forward and reverse Bedgraph files corresponding to the same sample were combined together and normalized to 1 million tags. In case of conflicting annotations, a single annotation was chosen according to the following hierarchy: intergenic < antisense < intron < exon < 5'UTR < proximal promoter. Differentially expressed TCS were called using the DESeq2 package.

Polymerase II EP and nascent RNA purification. Five grams of seedlings from a line where a NRPB2-FLAG construct covers a lethal npr2-1 mutation (described in ref.29) were grind to a fine powder. NRPB2 encodes the second largest subunit of the RNA polymerase II complex and since the FLAG-tagged protein covers a lethal mutation it ensures that all RNAPII complexes in the line contain a tagged NRPB2. RNA that was extracted by the miRNeasy Mini Kit (Qiagen, Germany) according to manufacturer instructions. Following purification, mRNA was fractionated into coding and noncoding fractions by using a library of 1000 UME reads (max_CFR=0.5, UME_max=100) per library. The UME (Universal Molecular Identifier) was identified as the most conserved group of primers that multiple regulatory pathways are activated during cold acclimation in Arabidopsis, delays leaf senescence and extends plant longevity. The UME primer set includes 2 µg of an antibody (Anti-Total RNA polymerase II 8WG16, ab817, abcam, UK) were added to solubilized chromatin. Quantitative analysis was performed on captured DNA by qPCR (Bio-Rad). See Supplementary Data 4 for oligonucleotide primers. qPCR enrichments were calculated as the ratio of product of interest from IP sample to the corresponding input sample. Error bars represent standard error of the mean resulting from at least three independent replicates. Source data can be found in the Source Data file.

Code availability. All supplementary code for the data analysis pipelines described is available at [https://github.com/Maxim-Ivanov/Kindgren_et_al_2018]. The authors declare that all other data supporting the findings of this study are available from the corresponding author upon request. The source data underlying all figures can be found in the Source Data file. A reporting summary for this Article is available as a Supplementary Information File.

Received: 13 April 2018 Accepted: 10 October 2018
Published online: 01 November 2018

References

1. Jensen, TorbenH., Jacquier, A. & Libri, D. Dealing with pervasive transcription. Mol Cell 52, 473–484 (2013).
2. Liu, F., Marquardt, S., Lister, C., Swiezewski, S. & Dean, C. Targeted 3' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. Science 327, 94–97 (2010).
3. Seo, J. S. et al. ELF18-INDUCED LONG-NONCODING RNA associates with mediator to enhance expression of innate immune response genes in Arabidopsis. Plant Cell 29, 1024–1038 (2017).
4. Fedak, H. et al. Control of seed dormancy in Arabidopsis by a cis-acting noncoding antisense transcript. Proc. Natl Acad. Sci. 113, E7846–E7855 (2016).
5. Mato, A. et al. Arabidopsis transcriptome analysis under drought, cold, high-salinity andABA treatment conditions using a tiling array. Plant Cell Physiol. 49, 1135–1149 (2008).
6. Liu, J. et al. Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. Plant Cell 24, 4333–4345 (2012).
7. Deveson, I. W., Hardwick, S. A., Mercer, T. R. & Mattick, J. S. The dimensions, dynamics, and relevance of the mammalian noncoding transcriptome. Trends Genet. 33, 464–478 (2017).
8. Ard, R., Allshire, R. C. & Marquardt, S. Emerging properties and functional consequences of noncoding transcription. Genetics 207, 357–367 (2017).
9. Djebali, S. et al. Landscape of transcription in human cells. Nature 489, 101–108 (2013).
10. Medina, Jn, Bargues, M., Terol, J., Pérez-Alonso, M. & Salinas, J. The Arabidopsis CBF gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration. Plant Physiol. 119, 463–470 (1999).
11. Medina, J., Catalá, R. & Salinas, J. The CBFs: Three Arabidopsis transcription factors to cold acclimate. Plant Sci. 180, 3–11 (2011).
12. Gusta, L. V., Wissiewski, M. E. & Tanino, K. K. Plant Cold Hardiness: From the Laboratory to the Field. (CABI, 2009).
13. Fowler, S. & Thomashow, M. F. Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell 14, 1675–1690 (2002).
14. Zhao, C. et al. Mutational evidence for the critical role of CBF transcription factors in cold acclimation in Arabidopsis. Plant Physiol. 171, 2744–2759 (2016).
15. Maruyama, K. et al. Identification of cold-inducible downstream genes of the Arabidopsis DREB1A/CBF3 transcriptional factor using two microarray systems. Plant J. 58, 982–993 (2004).
16. Kasuga, M., Liu, Q., Misura, S., Yamaguchi-Shinozaki, K. & Shinozaki, K. Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat. Biotechnol. 17, 287 (1999).
17. Sharabi-Schwager, M., Lers, A., Samach, A., Guy, C. L. & Porat, R. Overexpression of the CBF2 transcriptional activator in Arabidopsis delays leaf senescence and extends plant longevity. J. Exp. Bot. 61, 261–273 (2010).
18. Ioao-Ottosen, K. R., Gilmore, S. I., Zarka, D. G., Schabenberger, O. & Thomashow, M. F. Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. Science 280, 104–106 (1998).
19. Gilmore, S. J., Seblot, A. M., Salazar, M. P., Everard, J. D. & Thomashow, M. F. Overexpression of the Arabidopsis CBF3 transcriptional activator mimics
multiple biochemical changes associated with cold acclimation. Plant Physiol. 124, 1854–1865 (2014).

20. Gilmour, S. J., Fowler, S. G. & Thomashow, M. F. Arabidopsis transcriptional activators CBF1, CBF2, and CBF3 have matching functional activities. Plant Mol. Biol. 54, 767–781 (2004).

21. Swiezewski, S., Liu, F., Magusin, A. & Dean, C. Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycistron target. Nature 462, 799–802 (2009).

22. Kim, D.-H. & Sung, S. Vernalization-triggered intragenic chromatin-loop formation by long noncoding RNAs. Dev. Cell. 40, 302–312.e4 (2017).

23. Whittaker, C. & Dean, C. The FLC locus: a platform for discoveries in epigenetics and adaptation. Annu. Rev. Cell. Dev. Biol. 33, 555–575 (2017).

24. Pelechano, V., Wei, W. & Steinmetz, L. M. Genome-wide quantification of 5′-phosphorylated mRNA degradation intermediates for analysis of ribosome dynamics. Nat. Protoc. 11, 359 (2016).

25. Mathias Nielsen, Ryan Ard, Xueyuan Leng, Maxim Ivanov, Peter Kidgern, Vicent Pelechano, Sebastian Marquardt. Transcription-driven chromatin repression of intragenic promoters. Preprint at bioRxiv https://doi.org/10.1101/797144 (2018).

26. Zarka, D. G., Vogel, J. T., Cook, D. & Thomashow, M. F. Cold induction of Arabidopsis CBF genes involves multiple ICE (inducer of CBF expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. Plant Physiol. 133, 910–918 (2003).

27. Alonso, J. M. et al. Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science 301, 653–657 (2003).

28. Thalhammer, A., Hincha, D. K. & Zuther, E. in Plant Cold Acclimation: Methods and Protocols (eds Hincha, D. K. & Zuther, E.) 15–24 (Springer New York, New York, NY, 2014).

29. Agarwal, M. et al. A RiR3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. J. Biol. Chem. 281, 37637–37645 (2006).

30. Kuribara, Y. et al. Surveillance of 3′ non-coding transcripts requires FRI1 and XRN3 in Arabidopsis. G3 (Bethesda) 2, 487–498 (2012).

31. Lange, H. et al. The RNA helicases AtMTR4 and HEN2 target specific subsets of nuclear transcripts for degradation by the nuclear exosome in Arabidopsis thaliana. PLoS Genet. 10, e1004564 (2014).

32. Hématy, K. et al. The zinc-finger protein SOPI1 is required for a subset of the nuclear exosome functions in Arabidopsis. PLoS Genet. 12, e1005817 (2016).

33. Sikorski, P. J. et al. Distinct 18S RNA precursors are targets of the exosome complex, the exoribonuclease RRP6L2 and the terminal nucleotidylytransferase TRL1 in Arabidopsis thaliana. Plant J. 83, 991–1004 (2015).

34. Proudfoot, N. J. Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. Science 352, aad9926 (2016).

35. Rosonina, E., Kaneko, S. & Manley, J. L. Terminating the transcript: breaking up is hard to do. Genes Dev. 20, 1050–1056 (2006).

36. Krzyztop, M. et al. Defective XRN3-mediated transcription termination in Arabidopsis affects the expression of protein-coding genes. Plant J. 93, 1017–1031 (2018).

37. Ghildiyal, M. & Zamoore, P. D. Small silencing RNAs: an expanding universe. Nat. Rev. Genet. 10, 94 (2009).

38. Hobson, David. et al. RNA Polymerase II Collission Interrupts Converging Transcription. Mol. Cell 48, 365–374 (2012).

39. Krells, J. A. et al. Transcription changes for Arabidopsis in response to salt, osmotic, and cold stress. Plant Physiol. 130, 2129–2141 (2002).

40. DebRoy, S. et al. A riboswitch-containing sRNA controls gene expression by sequestration of a response regulator. Science 345, 937–940 (2014).

41. Mellin, J. R. et al. Sequestration of a two-component response regulator by a riboswitch-regulated noncoding RNA. Science 345, 940–943 (2014).

42. Hongay, C. F., Grisa, F., P. L., Galitski, T. & Fink, G. R. Antisense transcription of the nuclear exosome functions in Arabidopsis. Genetics 180, 207–218 (2008).

43. Marquardt, S. et al. Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. Mol. Cell 54, 156–165 (2014).

Acknowledgements

We would like to thank Prof. Åsa Strand and Sofie Grönlund for assistance with the freezing test. We thank Prof. Peter Brodersen for sharing seed and assistance with the Luciferase assay. We thank Prof. Vicent Pelechano for help with the 5′CAP-seq. We thank Janum Dilgen for technical assistance, Jan Hostrup for plant care and members of the S.M. laboratory for critical reading of the manuscript. Research in the laboratory of S. M. is supported by a Hallas-Møller Investigator award by the Novo Nordisk Foundation NNF15OC0014202, a Copenhagen Plant Science Centre Young Investigator Starting Grant. In addition, this project has received funding from the European Research Council (ERC) and the Marie Curie Actions under the European Union’s Horizon 2020 research and innovation programme STG2017-757411 (S.M.) and MSCA IF 703085 (P.K.). R.A. is supported by an EMBO LTF (ALTF 463–2016).

Author contributions

P.K., R.A., and M.I. performed the experiments. P.K., R.A., and M.I. analyzed the data. S. M. supervised the project. P.K. and S.M. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07010-6.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.