

# Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target

Szymon Swiezewski<sup>1</sup>, Fuquan Liu<sup>1</sup>, Andreas Magusin<sup>1</sup> & Caroline Dean<sup>1</sup>

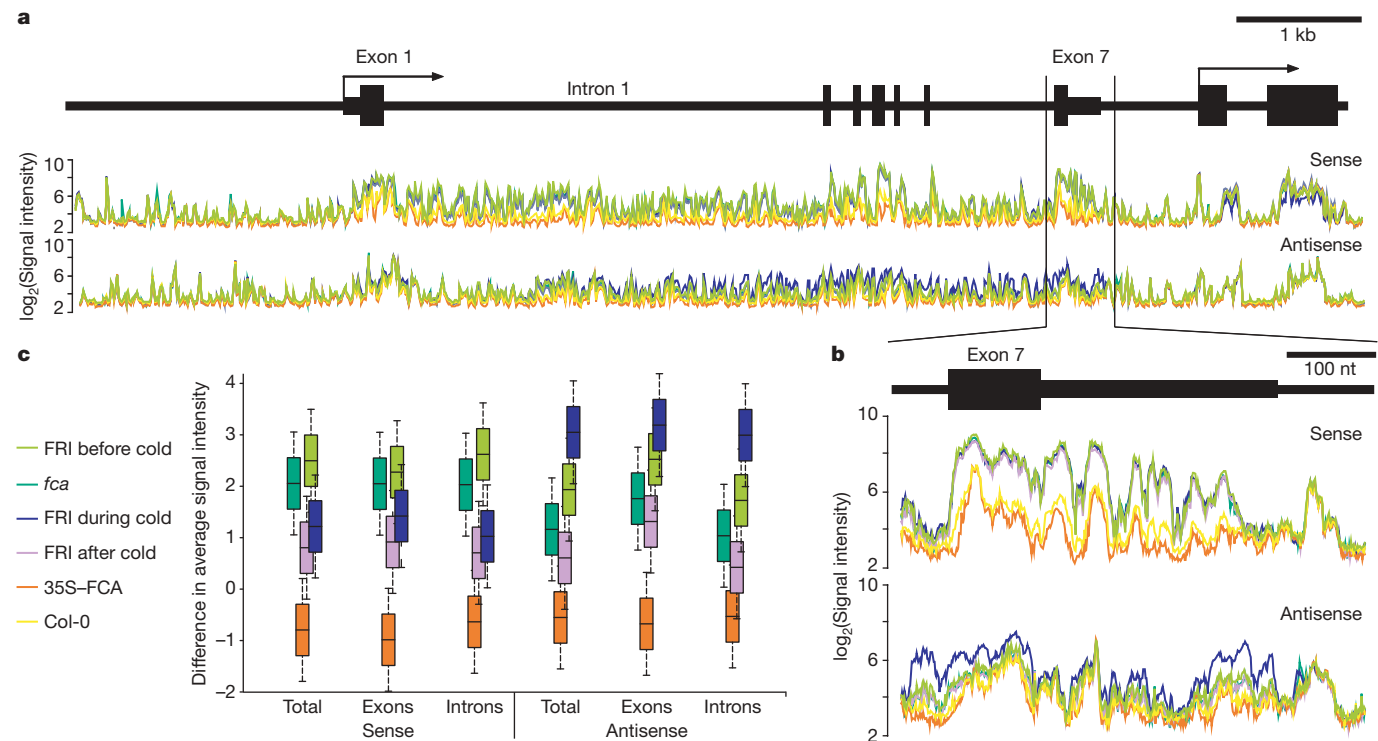
Transcription in eukaryotic genomes generates an extensive array of non-protein-coding RNA, the functional significance of which is mostly unknown<sup>1</sup>. We are investigating the link between non-coding RNA and chromatin regulation through analysis of *FLC*—a regulator of flowering time in *Arabidopsis* and a target of several chromatin pathways. Here we use an unbiased strategy to characterize non-coding transcripts of *FLC* and show that sense/antisense transcript levels correlate in a range of mutants and treatments, but change independently in cold-treated plants. Prolonged cold epigenetically silences *FLC* in a Polycomb-mediated process called vernalization<sup>2</sup>. Our data indicate that upregulation of long non-coding antisense transcripts covering the entire *FLC* locus may be part of the cold-sensing mechanism. Induction of these antisense transcripts occurs earlier than, and is independent of, other vernalization markers<sup>3</sup> and coincides with a reduction in sense transcription. We show that addition of the *FLC* antisense promoter sequences to a reporter gene is sufficient to confer cold-induced silencing of the reporter. Our data indicate that cold-induced *FLC* antisense transcripts have an early role in the epigenetic silencing of *FLC*, acting to silence *FLC* transcription transiently. Recruitment of the Polycomb machinery then confers the epigenetic memory. Antisense transcription events originating from 3' ends of genes might be a general mechanism to regulate the corresponding sense transcription in a condition/stage-dependent manner.

Advances in sequencing technology have enabled the discovery of many long non-coding transcripts. Their functions and evolution are poorly understood<sup>1</sup>, but many seem to be antisense transcripts that initiate and terminate near the terminators and promoters, respectively, of sense transcripts<sup>4</sup>. We have previously found non-coding transcripts originating from the 3' region of the gene encoding the floral repressor FLOWERING LOCUS C (*FLC*)<sup>5,6</sup>. *FLC* is upregulated by FRIGIDA and transcriptionally silenced through the activity of an RNA recognition motif (RRM) protein (*FCA*) and the K4 histone demethylase activity of FLOWERING LOCUS D (*FLD*)<sup>5,6</sup>. *FLC* is also repressed by prolonged cold and epigenetically silenced by a Polycomb mechanism involving a PHD-PRC2 (plant homeodomain-Polycomb repressive complex 2) complex<sup>2</sup>, in a process known as vernalization<sup>3,7</sup>. To understand the roles of non-coding RNA in these different types of *FLC* regulation we undertook an unbiased analysis of *FLC* transcripts in a selection of genotypes and environments. We designed a single nucleotide resolution array (NimbleGen/Affymetrix platform) of 25-base oligonucleotides covering both strands of *FLC* and the surrounding 50 kb region. We hybridized the arrays with total RNA (ribosomal RNA reduced through the use of Invitrogen RiboMinus—see Supplementary Methods) isolated from the following genotypes and treatments; Columbia (wild type), *FRIGIDA*+ (*FRI*), *fca*, 35S::*FCA* $\gamma$  (an overexpression of *FCA*), *FRI* seedlings that had been cold-treated for

14 days, *FRI* seedlings that had been cold-treated then followed by 7 days of further growth at 20 °C. Fourteen days of cold was chosen with the aim of identifying the early steps in vernalization; this is insufficient to saturate the vernalization requirement but does result in stable reduction in *FLC* expression.

Despite considerable differences in hybridization efficiencies between the different oligonucleotides there was excellent reproducibility between biological replicates (Fig. 1a, b). Comparison of the hybridization profile between the different genotypes and environments confirmed all previous findings of FRIGIDA and *FCA* function on *FLC* expression<sup>8,9</sup>, so validating the approach. Means of normalized hybridization data for each oligonucleotide in each treatment are shown in Figure 1a, b (*FLC*) and Supplementary Fig. 1 (whole 50 kb). Extensive *FLC* antisense transcripts were detected in all genotypes and treatments with levels correlating positively with sense transcript levels across all but one of the comparisons (Fig. 1). Additional quantitative PCR with reverse transcription (qRT-PCR) analysis of sense/antisense levels corroborated the array data and analysis of additional autonomous pathway mutants confirmed the close connection between sense and antisense levels (Supplementary Fig. 2e). This is consistent with the observation from many genome-wide studies showing a positive correlation in expression of most sense/antisense pairs<sup>10,11</sup>. The exception was cold-treated seedlings in which the ratio of sense to antisense *FLC* transcript changed significantly; antisense transcript accumulated over the majority of *FLC*, extending beyond the *FLC* 5' start and 3' polyadenylation sites, whereas levels of sense mRNA did not change significantly (Fig. 2a). Cold-induced accumulation of antisense transcripts for the other ten genes on the array was not observed (Supplementary Fig. 1). The accumulation of antisense *FLC* RNA was transient; levels returned to almost non-vernalized levels in plants returned to the warm for 7 days following 14 days cold (Fig. 1c). To determine the structure of the antisense transcripts we performed RACE (rapid amplification of cDNA ends) experiments. These defined several, cold-induced antisense *FLC* transcripts, with heterogeneous 5' ends initiating in a 100-nucleotide region immediately 3' to the sense *FLC* polyadenylation site (Fig. 2b, Supplementary Fig. 2). This coincides with the localized small RNA-induced heterochromatic region at the 3' end of *FLC* we have previously described<sup>5</sup>. As a parallel to HOTAIR (Hox antisense intergenic RNA)<sup>12</sup> we have called *FLC* antisense transcripts initiating from this region COOLAIR (cold induced long antisense intragenic RNA). RACE experiments showed that a proportion of the transcripts are either capped or carry a 5' triphosphate group, have long (60–100 adenosines) poly-A tails and are spliced (Supplementary Fig. 2b, c). We find alternatively processed forms, with respect to both polyadenylation site and splicing. Our data indicate that these antisense transcripts are produced from a DNA template as the splice sites are canonical and do not overlap with sense transcript splice sites. The antisense *FLC* transcripts also

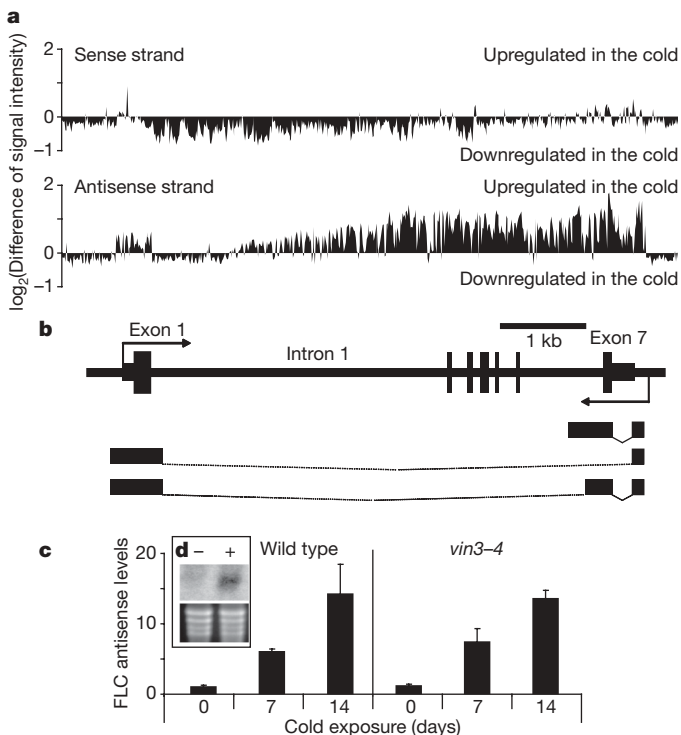
<sup>1</sup>John Innes Centre, Colney, Norwich, NR4 7UH, UK.



**Figure 1** | *FLC* is associated with extensive antisense transcription.

**a**, Signal intensity ( $\log_2$ ) for the median of a 10-base-pairs staggered window along the genomic region of *Arabidopsis* chromosome V containing *FLC*. Genomic structure of *FLC*; black boxes represent exons; lines represent IGR and introns. **b**, Magnified region of *FLC* 3' end showing *FLC* antisense transcript start. **c**, Boxplots illustrating the difference in average signal

intensity for sense and antisense strands compared to Col-0 (Columbia wild type). The central box represents the data between the quartiles, the median is shown by the line through the centre of box; and the whiskers extend out to the extremes of the data. Total, whole *FLC* gene; exons, sense *FLC* exons, introns, sense *FLC* introns.

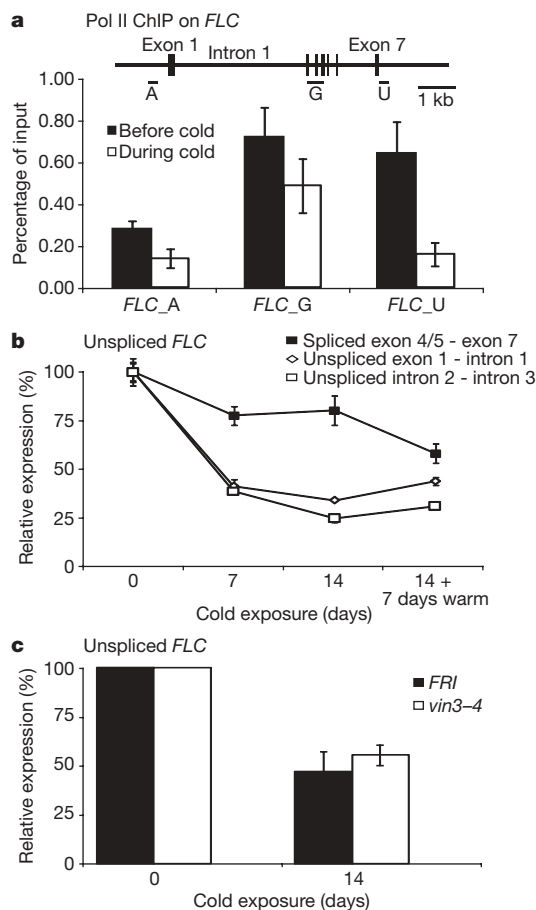


**Figure 2** | *FLC* antisense transcripts are strongly induced by cold.

**a**, Difference of signal intensities ( $\log_2$ ) for cold-treated and untreated plants along sense and antisense strands of *FLC*. **b**, *FLC* region as in Fig. 1a with antisense transcript structure shown below. **c**, qPCR quantification of relative *FLC* antisense transcript levels in wild type and *vin3-4*. Data are means and s.e.m. ( $n = 2$ ). **d**, Northern blot of *FLC* antisense transcripts in control seedlings (–) and those cold-treated for 14 days (+).

extend beyond each end of the sense transcript (Fig. 2b, Supplementary Fig. 2a,d).

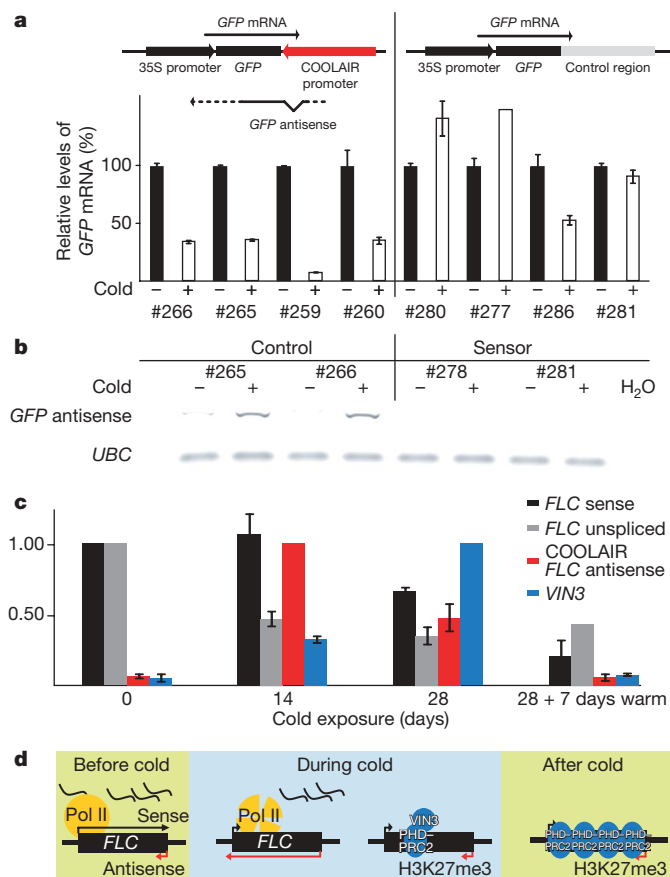
Given that Polycomb silencing of *FLC* is cold-dependent we explored the timing of COOLAIR induction relative to other known markers in the vernalization process. The earliest reported step to date is the cold-induction of *VIN3* (VERNALIZATION INSENSITIVE 3), a PHD protein required for vernalization<sup>3</sup>. *VIN3* expression is first detectable 20 days after transfer to cold, is maximal 40 days after transfer and is then undetectable 3 days after return to warm conditions<sup>3</sup>. *VIN3* heterodimerizes with a related PHD protein *VRN5* (VERNALIZATION 5, ref. 13) and induces formation of a PHD–PRC2 complex that associates with a specific domain in the first intron of *FLC* after prolonged cold<sup>2</sup>. This association triggers spreading of *VRN5* along *FLC* and significant enhancement of histone H3 Lys 27 trimethylation (H3K27me3) levels throughout the locus, to levels required for maintenance of the silencing during subsequent growth of the plant<sup>2</sup>. COOLAIR induction was independent of *VIN3* (Fig. 2c) and *VIN3* induction was found to be much later than maximal COOLAIR induction (Fig. 4c). Because COOLAIR induction might precede the nucleation of the Polycomb silencing at *FLC*, we analysed the timing of repression of *FLC* transcription by cold. This was achieved in three ways; analysis of the exon/intron hybridization to the array, through quantitative PCR analysis of nascent and spliced sense *FLC* transcripts, and using chromatin immunoprecipitation experiments with antibodies that react with all forms of RNA polymerase II (RNA polII) (Fig. 1c, 3a, 3b). Using all three assays *FLC* sense transcription was found to decrease much faster than spliced sense transcript after the plants were transferred to cold (Fig. 3b), and this was again independent of *VIN3* (Fig. 3c). Thus, an early step in vernalization seems to be cold-induced upregulation of antisense transcription linked to sense transcriptional repression that is independent of, and precedes, the maximum induction of *VIN3*. This is reminiscent of *bx1* silencing where non-coding RNA transcription has been shown to be involved in its 'early' repression



**Figure 3 | Vernalization-induced transcriptional repression of *FLC*.** **a**, ChIP using RNA polII-specific antibody. Regions of *FLC* monitored in the ChIP are shown as bars below a schematic of the genomic structure of *FLC*. Controls for the ChIP are in Supplementary Information. Data are means and s.e.m. ( $n = 3$ ). **b**, qPCR quantification of *FLC* mRNA and *FLC* unspliced sense transcript after different cold treatments. Data are means and s.e.m. ( $n = 2$ ). **c**, qPCR quantification of *FLC* unspliced sense transcript showing cold-induced downregulation is unaffected in *vin3-4*. Data are means and s.e.m. ( $n = 2$ ).

before Polycomb action<sup>14</sup>. The slower reduction of *FLC* spliced sense transcript may indicate a stable pool of reserve transcript maintained after transfer to cold that decreases after several weeks of cold, coincident with the decrease in antisense transcription (Fig. 4c).

*FLC* sense transcriptional silencing coincided with maximal COOLAIR expression so we tested whether COOLAIR was sufficient to cause transcriptional silencing. The COOLAIR promoter (aka *FLC* sense 3' region) was fused downstream of the coding region of a 35S (strong constitutive) promoter-green fluorescent protein (GFP) fusion (Fig. 4a). The 3' region of a gene showing no antisense transcription (*rbcS3A*) was used as control. There was no consistent change in expression of the sensor carrying the *rbcS3A* 3' region; whereas the COOLAIR promoter was found to reproducibly confer cold-dependent silencing of the sensor construct (a representative set from a larger number of transformants is shown in Fig. 4a). This was associated with cold-induction of an antisense transcript to the GFP sensor construct, showed by strand-specific RT-PCR analysis (Fig. 4b). Sequencing of the PCR products showed the transcript to be spliced at the same position as COOLAIR (Supplementary Fig. 2d) and extend from the COOLAIR promoter to at least half way through the transgene (Fig. 4a). The cold-induced silencing of the GFP sensor construct was transient and expression increased once plants were returned to warm conditions (Supplementary Fig. 3). This reinforces the view that it is the process of transcription rather than specific sequences within the *FLC* antisense transcripts that cause the cold-induced silencing. Supporting this view, we have found the



**Figure 4 | *FLC* antisense promoter region is sufficient to confer cold-dependent silencing of sense transcription.** **a**, Schematic representation and expression of GFP sensor lines. Warm-grown (black) or 14 days cold-treated (white). Structure of GFP antisense transcript shown below. **b**, GFP antisense transcript levels in GFP sensor lines. GFP sense expression data for line #278 shown in Supplementary Fig. 4. UBC, *Arabidopsis* ubiquitin-conjugating enzyme gene. **c**, Relative temporal expression of spliced and unspliced *FLC* sense transcript, COOLAIR and *VIN3* before cold, after 14 days and 28 days of cold and then 7 days after transfer back to warm. Data are means and s.e.m. ( $n = 2$ ). **d**, Model indicating sequence of events. Green/blue, warm/cold, respectively.

COOLAIR promoter alone can drive cold-induced expression of LUC sequences (Supplementary Fig. 4). In the context of the endogenous *FLC* gene, the cold-induced increase in the antisense transcripts is linked to reduction in levels of nascent sense transcript and reduction in occupancy of RNA polII at *FLC*. This, together with the fact that COOLAIR transcripts extend beyond the transcription start site of *FLC*, indicates that antisense transcription might contribute to *FLC* transcriptional repression through promoter interference (Fig. 4d).

A similar cold-induced transcriptional silencing has been described for *PHO84* in *S. cerevisiae*<sup>15</sup>. Cold-induced exosome-dependent stabilization of antisense transcripts led to histone deacetylase recruitment and stable *PHO84* silencing. Although we cannot exclude a cold-dependent post-transcriptional stabilization of COOLAIR, specific sequences or RNA structures seem to be dispensable as the promoter region alone is sufficient for cold-induced transcript production. However, we also see a consistent cold-dependent decrease in RNA polII occupancy at the COOLAIR promoter and elsewhere in *FLC*. This could indicate involvement of post-transcriptional mechanisms in COOLAIR induction or involvement of an alternative DNA-dependent RNA polymerase (although induction still occurs in a polIV/polV double mutant<sup>16</sup>; Supplementary Fig. 5). It may also simply reflect that the reduction in sense transcription is masking any increase in RNA polII occupancy associated with increased antisense transcription.

In summary, our study has demonstrated the environmental induction of long, non-coding, antisense transcripts covering an entire gene. These were found to be sufficient to induce silencing of a linked reporter. In the context of endogenous *FLC* silencing we propose that COOLAIR is involved in early, cold-dependent and transient, transcriptional silencing of *FLC*. This silencing would then be reinforced and epigenetically maintained by the Polycomb machinery. A possible sequence of events during the cold might be COOLAIR induction, suppression of sense transcription, VIN3 upregulation and PHD–PRC2 complex nucleation<sup>2</sup>. In agreement with this we find COOLAIR expression continues longer in the cold in the absence of VIN3, supporting a model where the Polycomb machinery acts after COOLAIR and silences the whole *FLC* locus preventing further sense and antisense transcription (Supplementary Fig. 6). Recent data on PRC2 complex interaction with non-coding RNA transcripts<sup>12,17</sup> might indicate that COOLAIR has additional, Polycomb-dependent functions in vernalization. Cold-responsiveness of an *FLC*–*GUS* transgene lacking the COOLAIR sequences and a possible Polycomb-independent silencing pathway that might operate *in trans*<sup>18</sup> may also indicate mechanistic redundancy between different vernalization pathways. Our data are consistent with COOLAIR transcription being sufficient to silence linked sequences in a Polycomb-independent manner; however, these non-coding RNAs may function in both *cis* and *trans*<sup>19</sup>.

In a wider context, our discoveries on *FLC* may have broad relevance to other organisms where a genome-wide association of non-coding RNA transcript start sites with 3' regions has been observed<sup>4</sup>. Many studies have found a strong correlation between sense and antisense transcript abundance<sup>10,11</sup>. On the basis of our data we speculate that many antisense transcription events originating from the 3' end of genes might regulate the sense partner in a condition/stage dependent manner. Continuous production of low levels of non-coding antisense transcripts maybe the cost of this regulatory mechanism.

## METHODS SUMMARY

**Plant growth conditions and vernalization treatment.** All genotypes used have been described previously<sup>8,16,20</sup>. All seeds were surface-sterilized and grown on MS medium minus glucose. The 14-day cold-treatment was carried out on plants pre-grown for 7 days at 20 °C. Cold treatment was at ~5 °C under short-day conditions, 8 h white light (10 mmol photons m<sup>-2</sup> s<sup>-1</sup>), 16 h dark. Seedlings were then either harvested directly or transferred to long-day conditions of 16 h white light (57 mmol photons m<sup>-2</sup> s<sup>-1</sup>), 8 h dark and grown for 7 days at 20 °C. Non-vernalized seedlings were cold-treated for 2 days, then transferred to long-day conditions for 14 days.

**Array.** Total RNA that was rRNA-depleted was labelled and hybridized by Cogenics on custom-design Affymetrix/Nimble Express GeneChips. For description of array design and analysis see Methods. COOLAIR sequences have been deposited in GenBank under accession numbers GQ352646 and GQ342259. Array data have been deposited to the Gene Expression Omnibus repository under the GEO ID GSE16977.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 16 July; accepted 30 October 2009.

1. Ponting, C. P., Oliver, P. L. & Reik, W. Evolution and functions of long noncoding RNAs. *Cell* **136**, 629–641 (2009).

2. De Lucia, F. *et al.* A PHD–Polycomb repressive complex 2 triggers the epigenetic silencing of *FLC* during vernalization. *Proc. Natl Acad. Sci. USA* **105**, 16831–16836 (2008).
3. Sung, S. & Amasino, R. M. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature* **427**, 159–164 (2004).
4. He, Y. *et al.* The antisense transcriptomes of human cells. *Science* **322**, 1855–1857 (2008).
5. Swiezewski, S. *et al.* Small RNA-mediated chromatin silencing directed to the 3' region of the *Arabidopsis* gene encoding the developmental regulator, *FLC*. *Proc. Natl Acad. Sci. USA* **104**, 3633–3638 (2007).
6. Liu, F. *et al.* The *Arabidopsis* RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate *FLC*. *Mol. Cell* **28**, 398–407 (2007).
7. Bastow, R. *et al.* Vernalization requires epigenetic silencing of *FLC* by histone methylation. *Nature* **427**, 164–167 (2004).
8. Lee, I., Michaels, S. D., Masshardt, A. S. & Amasino, R. M. The late-flowering phenotype of *FRIGIDA* and mutations in *LUMINIDEPENDENS* is suppressed in the Landsberg *erecta* strain of *Arabidopsis*. *Plant J.* **6**, 903–909 (1994).
9. Macknight, R. *et al.* FCA, a gene controlling flowering time in *Arabidopsis*, encodes a protein containing RNA-binding domains. *Cell* **89**, 737–745 (1997).
10. Oeder, S., Mages, J., Flicek, P. & Lang, R. Uncovering information on expression of natural antisense transcripts in Affymetrix MOE430 datasets. *BMC Genomics* **8**, 200 (2007).
11. RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group) and the FANTOM Consortium. Antisense transcription in the mammalian transcriptome. *Science* **309**, 1564–1566 (2005).
12. Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell* **129**, 1311–1323 (2007).
13. Greb, T. *et al.* The PHD finger protein VRN5 functions in the epigenetic silencing of *Arabidopsis FLC*. *Curr. Biol.* **17**, 73–78 (2007).
14. Petruk, S. *et al.* Transcription of *bx*d noncoding RNAs promoted by trithorax represses *Ubx* in *cis* by transcriptional interference. *Cell* **127**, 1209–1221 (2006).
15. Camblong, J. *et al.* Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* **131**, 706–717 (2007).
16. Onodera, Y. *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622 (2005).
17. Pandey, R. R. *et al.* *Kcnq1ot1* antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell* **32**, 232–246 (2008).
18. Sheldon, C. C., Finnegan, J. E. & Peacock, W. J. and Dennis, E. S., Mechanisms of gene repression by vernalization in *Arabidopsis*. *Plant J.* **59**, 488–498 (2009).
19. Camblong, J. *et al.* Trans-acting antisense RNAs mediate transcriptional gene cosuppression in *S. cerevisiae*. *Genes Dev.* **23**, 1534–1545 (2009).
20. Sung, S., Schmitz, R. J. & Amasino, R. M. A PHD finger protein involved in both the vernalization and photoperiod pathways in *Arabidopsis*. *Genes Dev.* **20**, 3244–3248 (2006).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank all members of the Dean laboratory for discussions, J. Ecker and B. Gregory (Salk Institute) for help with the RiboMinus technique, C. Lister for help with cloning steps, M. Lenhard and H. Breuninger for seeds of LUC control plants. This research was supported by the BBSRC Core Strategic Grant to the John Innes Centre and BBSRC BB/D010799/1, NERC NE/C507629 and EU SIROCCO LSHG-CT-2006-037900 grants to C.D.

**Author Contributions** C.D. and S.S. designed the study. S.S. performed most experiments. S.S. and F.L. performed the array experiment. A.M. and S.S. analysed microarray data. C.D. and S.S. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). Correspondence and requests for materials should be addressed to C.D. ([caroline.dean@bbsrc.ac.uk](mailto:caroline.dean@bbsrc.ac.uk)).

## METHODS

**RNA quantification.** For the *FLC* antisense northern blot a probe was generated using PCR on genomic DNA with the primers FLCexon7 5'-GGAGATAATCA TCATGTGGGAGCA-3' and FLC3UTR 5'-CTCACACGAATAAGGTACAAA GTTC-3'. Labelling was performed with FLCexon7 primer in a primer extension reaction.

*FLC* antisense quantification was undertaken using qPCR with primers F 5'-TTTTTTTTTTTTTACTGCTTCCA-3' and R 5'-CACACCACCAAATA ACAACCA-3' known as set 1 and normalized against the *Arabidopsis UBC* gene<sup>21</sup>. The first strand cDNA was synthesized with an oligo(dT) primer.

*FLC* sense spliced (spliced exon 4/5, 5'-AGCCAAGAAGACCCGAACCTCA-3'; exon 7, 5'-TTTGTCAGCAGGTGACATC-3') and *FLC* unspliced (unspliced exon 1, 5'-TTCTCTCCGGCGATAAGTA-3'; unspliced intron1, 5'-TCACT CAACAACATCGAGCAC-3'; unspliced intron2, 5'-CGCAATTTTCATA GCCCTTG-3'; unspliced intron3, 5'-CTTTGTAATCAAAGGTGGAGAGC-3') quantification was achieved using a mixture of gene-specific-primer-primed cDNA (exon 7, unspliced intron1, unspliced intron3, UBC\_R).

*GFP* mRNA quantification was achieved using the primers GFP\_qPCR\_LP, 5'-TGCAGTGCTTCTCCCGTTAC-3'; GFP\_qPCR\_RP, 5'-GGTCTCTCCTG CACGTATC-3'. The *GFP* antisense transcript was detected by RT-PCR with primers GFPas2\_F, 5'-TCAAGGACGACGGGAACACTAC-3' (or GFPas1\_F, 5'-GTCCACACAATCTGCCCTT-3') and R, 5'-TTGACAGAAGTGAAGAACA CATAACA-3' on a cDNA synthesized with GFPas2\_F or GFPas1\_F using Phusion (Finnzymes).

*VIN3* levels were quantified using the primers VIN3\_F, 5'-GTATGG GATTGGGAGTGATGAT-3' and VIN3\_R, 5'-CAAAACAACCTGAAACCTG TGA-3'.

The *FLC* antisense RNA shown in Supplementary Fig. 2 was quantified by qPCR using primers F (5'-ACCTTATTCGTGTGAGAATTGC-3') and R (5'-TTGACAGAAGTGAAGAACAACATACA-3'), (known as set 6) on cDNA synthesized using oligo(dT) and set6\_F and normalized against *UBC*.

cDNA was synthesized using SuperScript III (Invitrogen) and analysed by quantitative PCR on an OPTICON2 instrument and using SYBR Green Jump-Start Taq ReadyMix from Sigma.

**Constructs.** pMDC43<sup>22</sup> and the Gateway system (Invitrogen) was used to clone an *FLC* antisense promoter region (5'-CACCTTCCGGTGACTCTCCCACTA-3' and 5'-TTTAACAACCTTTCACCTT-3') or the equivalent region from the *Arabidopsis rbc53A* gene (5'-CACCAACACCCGTCAGTCCAA-3' and 5'-GGTGGGATCCAAAATCAAG-3'). The resulting plasmids were transformed into *Agrobacterium tumefaciens* strain GV3101 and introduced into Col-0 FRI(Sf2) plants<sup>23</sup> by *Agrobacterium*-mediated floral dip transformation<sup>24</sup>.

The COOLAIR promoter fusion construct was constructed by MultiSite Gateway (Invitrogen)-mediated recombination into a pRK290-based *Agrobacterium* binary vector. The fragments were amplified using the primers COOLAIRprom\_F, 5'-GGGGACTGCTTTTTGTACAAACTTGCAATCTTC CGGTGACTCTCC-3' and COOLAIRprom\_R, 5'-GGGGACAACCTTTGTATA GAAAAGTTGTTTTAACAACCTTTCACCTT-3' on genomic DNA from Col-0 plants; BOX2\_F, 5'-GGGGACAAGTTTGTACAAAAAGCAGCGTA CGACGACTCGTCCGTCCTGTA-3' and BOX2\_R, 5'-GGGGACCACTTTGT ACAAGAAAGCTGGGTTTGATCAATCCACAGTTTTC-3' on pCambia1381Z<sup>25</sup>; LUC\_F, 5'-GGGGACAGCTTCTGTACAAAGTGGTGTAGATAGATAGCGATT CGGT-3' and LUC\_R, 5'-GGGGACAACCTTTGTATAATAAAGTTGGACCCGATC TAGTAACATAGA-3' on yy376 (ref. 26). The *FLC*-LUC translational fusion was described previously<sup>27</sup>. The control plasmid was created by fusing a 1-kilobase sequence upstream from the ATG codon of At3g63530, a gene which expression is not affected by cold, into a pBar vector<sup>28</sup>.

**RACE.** *FLC* antisense transcript mapping was performed as described in ref. 29 and the 5'RACE was performed using an Invitrogen RACE kit.

**Chromatin immunoprecipitation.** Chromatin immunoprecipitation and primers previously described in ref. 30 were used in combination with an RNA polIII antibody 8WG16 from Abcam (ab817). Actin was used as an internal control using primers described in ref. 30. The average precipitation for actin of the

non-vernallized samples was 2.69% of input, compared to 2.33% of input in vernallized samples.

**Microarray.** RNA was extracted using the method described in ref. 31. rRNA was removed with a RiboMinus Plant Kit (Invitrogen). The resulting RNA was labelled using the Affymetrix GeneChip Whole Transcript Sense Target Labelling Assay. Probes were labelled and hybridized by Cogenics on custom design Affymetrix/Nimble Express GeneChips. The arrays contained 25-nucleotide oligonucleotides tiled with single nucleotide resolution across both strands of the 50 kb region of chromosome 5 including *FLC*. We also printed probes corresponding to genes used for expression normalization corresponding to oligonucleotides used on the Affymetrix ATH1 array and shown by ref. 21 to be optimal as reference genes for transcript normalization. In addition, we included genes we had used previously to normalize expression in northern blots and qPCR experiments (250317at, 244918at, 259361at, 253287at, 259407at, 262909at, 253355at, 265256at). The microarrays were analysed using the statistical analysis language R<sup>32</sup>, using the LIMMA<sup>33</sup> and CNA<sup>34</sup> libraries under the BioConductor framework<sup>35</sup>. We used an algorithm that computes a global rank-invariant set of transcripts; which was then in turn employed to normalize all microarrays across experimental conditions<sup>35</sup>. Following normalization, linear models were fitted to the data according to ref. 33; and *t*-scores for all contrasts of interest were obtained. The *t*-scores were used as more reliable estimates of differential expression than averaged expression ratios as we let variation affect the quality of the scores. Segments of scores were obtained using the method of ref. 34; where *P*-values were obtained though permutations of *t*-scores contained in segments.

- Czechowski, T. *et al.* Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol.* **139**, 5–17 (2005).
- Curtis, M. D. & Grossniklaus, U. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* **133**, 462–469 (2003).
- Lee, I., Michaels, S. D., Masshardt, A. S. & Amasino, R. M. The late-flowering phenotype of *FRIGIDA* and mutations in *LUMINIDEPENDENS* is suppressed in the Landsberg erecta strain of *Arabidopsis*. *Plant J.* **6**, 903–909 (1994).
- Bechtold, N., Ellis, J. & Pelletier, G. In planta *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C.R. Acad. Sci. III* **316**, 1194–1199 (1993).
- Hajdukiewicz, P., Svab, Z. & Maliga, P. The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol. Biol.* **25**, 989–994 (1994).
- Yamamoto, Y. Y. *et al.* Gene trapping of the *Arabidopsis* genome with a firefly luciferase reporter. *Plant J.* **35**, 273–283 (2003).
- Mylne, J. S. *et al.* LHP1, the *Arabidopsis* homologue of *HETEROCHROMATIN PROTEIN1*, is required for epigenetic silencing of *FLC*. *Proc. Natl Acad. Sci. USA* **103**, 5012–5017 (2006).
- Völker, A., Stierhof, Y. D. & Jürgens, G. Cell cycle-independent expression of the *Arabidopsis* cytokinesis-specific syntaxin KNOLLE results in mistargeting to the plasma membrane and is not sufficient for cytokinesis. *J. Cell Sci.* **114**, 3001–3012 (2001).
- Liu, F. *et al.* The *Arabidopsis* RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate *FLC*. *Mol. Cell* **28**, 398–407 (2007).
- De Lucia, F. *et al.* A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of *FLC* during vernalization. *Proc. Natl Acad. Sci. USA* **105**, 16831–16836 (2008).
- Quesada, V., Macknight, R., Dean, C. & Simpson, G. G. Autoregulation of *FCA* pre-mRNA processing controls *Arabidopsis* flowering time. *EMBO J.* **22**, 3142–3152 (2003).
- R. Development Core Team. R: a language and environment for statistical computing. (<http://www.R-project.org>) (R Foundation for Statistical Computing, 2006).
- Smyth, G. K. in *Bioinformatics and Computational Biology Solutions using R and Bioconductor* (eds Gentleman, R., Carey, V. J., Huber, W., Irizarry, R. A. & Dudoit, S.) 397–420 (Springer, 2005).
- Olshen, A. B., Venkatraman, E. S., Lucito, R. & Wigler, M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Bioinformatics* **5**, 557–572 (2004).
- Pelz, C. R., Kulesz-Martin, M., Bagby, G. & Sears, R. C. Global rank-invariant set normalization (GRSN) to reduce systematic distortions in microarray data. *BMC Bioinformatics* **9**, 520 (2008).