Novel insights into gene silencing mechanisms in Zea mays and Arabidopsis thaliana

Hövel, I.

Publication date
2016

Document Version
Final published version

Chapter 7

Changes in chromosomal interaction patterns of *FLC* upon vernalization

Iris Hövel [a], Dimitris Zisis [b], Mariliis Tark-Dame [a], Rurika Oka [a], Pawel Krajewski [b], Maike Stam [a]

Author Affiliations:

[a] Swammerdam Institute for Life Sciences, University of Amsterdam, The Netherlands.

[b] Institute of Plant Genetics, Polish Academy of Sciences, Poznan, Poland

In preparation for submission

Abstract

In order to facilitate transcriptional regulation, chromatin in the interphase nucleus of eukaryotes is folded in a hierarchical manner. Thereby, specific chromatin domains can be formed, such as e.g. polycomb domains in *Drosophila melanogaster*. In these domains polycomb-silenced genes, marked with H3K27me3, cluster via long-range chromosomal interactions. Chromosomal interactions between particular polycomb targets have also been described for *Arabidopsis thaliana*. Here, we study the genome-wide chromatin interaction pattern of the *Flowering locus C* (*FLC*) gene, a polycomb target, in *A. thaliana*. *FLC* encodes a repressor of flowering and is targeted by polycomb silencing to allow flowering of *A. thaliana*. In *A. thaliana* winter-annual accessions, polycomb silencing of *FLC* is triggered by vernalization, a prolonged cold treatment. We examined the chromatin interaction pattern of *FLC* throughout the vernalization process to address the questions whether this pattern changes upon transcriptional repression and if any of the detected interacting sequences are co-regulated with *FLC*. We show that after vernalization, interaction frequencies increased between *FLC* and an H3K27me3 marked region ~ 55 kb downstream and decreased for regions with low H3K27me3 within a ~200 kb region around the *FLC* gene. These results indicate that, once the *FLC* gene itself is repressed by polycomb, it interacts preferably with other closely located polycomb targets rather than with non-polycomb targets. Furthermore, a general, transient increase of interactions with sequences on other chromosomes was noticed during the vernalization process. This increase is possibly caused by cold-stress induced chromatin decompaction that results in changes in the higher-order chromatin organization.

Introduction

Flowering in plants is well-controlled by transcriptional regulation. To allow flowering in *Arabidopsis thaliana*, the flowering repressor *Flowering locus C* (*FLC*) undergoes polycomb silencing. In winter-annual accessions, *FLC* silencing is induced by vernalization. Here we are interested in changes in chromosomal interactions of *FLC* during the vernalization process.

Chromatin is compactly organized in the interphase nucleus of all eukaryotes, and is organized in a hierarchical manner in order to facilitate transcription of desired DNA sequences. Chromosomes are generally not intermingled; each chromosome arm forms a separate, compact domain that has its distinct territory in the nucleus (Cremer and Cremer, 2001; Grob et al., 2014; Pombo and Branco, 2007; Berr and Schubert, 2007; Schubert et al., 2012). Chromatin is further organized into regions of heterochromatin, which is compact and associated with repressive histone marks, and euchromatin, which is less densely compacted and accessible for gene expression (Fransz and De Jong, 2011).

In mammals and Drosophila, chromosomes were found to be further organized in topological associated domains (TADs) of ~1 Mb or ~100 kb size, respectively (Sexton et al., 2012; Nora et al., 2012; Dixon et al., 2012). TADs span linear arrangements of co-expressed genes. Additionally, gene-rich and gene-poor TADs correlate with replication domains (Caron, 2013; Hiratani et al., 2008; Moindrot et al., 2012; Spellman and Rubin, 2002). TADs are dedicated to specific types of chromatin, active chromatin or repressed chromatin, such as polycomb-repressed chromatin or heterochromatin (Ciabrelli and Cavalli, 2015).

TADs have not been detected in all organisms studied thus far. For example, the model system for plants, *A. thaliana* seems to lack TADs (Feng et al., 2014; Grob et al., 2014; Wang et al., 2015). Not detecting TADs in *A. thaliana* could be attributed to the absence of linear arrangements of co-expressed genes, as well as its high gene density (Schmid et al., 2005; Swarbreck et al., 2008). Clustering of chromatin regions in *A. thaliana* has been found for pericentromeric heterochromatin, which is microscopically visible in the form of DNA-dense chromocenters (Fransz et al., 2002; Feng et al., 2014; Grob et al., 2014). Moreover, telomere regions were shown to cluster around the nucleolus and may interact with high frequencies *in cis* and *in trans* with each other (Feng et al., 2014; Grob et al., 2014). *In cis* interactions have also been observed for euchromatic regions enriched for the polycomb repressive complex 2 (PRC2) -dependent H3K27me3 mark (Feng et al., 2014; Wang et al., 2015). In polycomb mutants lacking H3K27me3 marks the frequency of such interactions decreased (Feng et al., 2014). These *in cis* interactions, however, are limited to a small number of genomic regions with extended enrichment for H3K27me3.

Several examples of regulatory chromosomal interactions have been described in great detail in a diversity of organisms. Chromosomal interactions may directly regulate gene expression, e.g. by enhancing transcription or by enforcing gene repression (Ciabrelli and Cavalli, 2015). The existence of distant enhancers that control gene expression by chromosomal interactions has been demonstrated for example by the locus control region regulating the beta-globin loci in mouse and human (Deng et al., 2012; Mitchell and Fraser, 2008; Palstra et al., 2008, 2003) or the hepta-repeat regulating the *booster 1* locus in maize (Belele et al., 2013; Louwers et al., 2009a; Stam et al., 2002). Also repressive chromosomal interactions can occure, for example in the form of polycomb repressive complex foci that act as domains for gene repression. In *Drosophila melanogaster* multiple scattered polycomb-targeted gene clusters have been shown to form exclusive long-range interactions with each other (Cléard et al., 2006; Lanzuolo et al., 2007; Tolhuis et al., 2011).

The *FLC* gene is coding for a MADS-box transcriptional repressor of flowering in *A. thaliana*. In winter-annual accessions, *FLC* expression is enhanced by its transcriptional regulator FRIGIDA (Zhu et al., 2015). To allow flowering, *FLC* is repressed by vernalization, a prolonged cold treatment. Extensive research on the transcriptional regulation of *FLC* revealed an intricate silencing process involving a specialized polycomb complex, non-coding transcripts as well as changes in chromosomal interactions. During vernalization, silencing of *FLC* is triggered by the assembly of a PHD and PRC2 protein complex on the locus (De Lucia et al., 2008). The resulting complex causes replacement of active H3K36me3 marks from the *FLC* region by repressive H3K27me3 marks (Yang et al., 2014). Vernalization induces a quantitative epigenetic memory for *FLC* repression, whereupon *FLC* progressively switches to a fully repressed state in individual cells (Angel et al., 2011). While nascent *FLC* transcript levels decrease during vernalization, the level of non-coding antisense transcripts, referred to as COOLAIR is upregulated (Swiezewski et al., 2009). COOLAIR enhances the down-regulation of *FLC*. However, disruption of COOLAIR expression does not disrupt vernalization (Helliwell et al., 2011). Another transcript, the sense strand, non-coding COLDAIR, is produced from *FLC* intron 1 upon cold induction (Heo and Sung, 2011). COLDAIR accumulates later in the silencing process than COOLAIR; it interacts with the histone methyltransferase subunit of PRC2 and is therefore thought to promote the association of the PRC2 complex to the *FLC* locus.

The active *FLC* gene forms a chromatin loop, indicated by interactions between its promotor and 3' region (Crevillén et al., 2013). Upon two weeks of vernalization the gene loop is disrupted. The loss of the gene loop is independent from the presence of polycomb effectors and is therefore interpreted as an early step in the stable repression of *FLC*. Life-cell imaging has revealed co-localization of *FLC* transgenes upon vernalization (Rosa et al., 2013), indicating a relation between repression and the physical clustering of *FLC* alleles. However, in a given nucleus one *FLC* allele can be active while the other is repressed (Berry et al., 2015). Both states are mitotically heritable, which indicates that the epigenetic

memory for *FLC* repression is stored in an allele-specific manner, hence *in cis* and not *in trans*.

Although the dynamics of the *FLC* gene loop structure during vernalization is known, the interaction dynamics of *FLC* with other sequences within the genome are yet to be defined. Here, we study the genome-wide chromatin interaction pattern of the *FLC* gene throughout the vernalization process. We hypothesized that the chromosomal interaction pattern of *FLC* changes with its epigenetic state throughout vernalization and that sequences interacting with *FLC* might be co-regulated with *FLC*. The interaction frequencies between the *FLC* gene body and other euchromatic sequences *in cis* or *in trans* was measured by using the Circular Chromosome Conformation Capture method followed by next generation sequencing (4C-Seq). When comparing the 4C interaction frequencies within a ~200 kb region around *FLC* for a non-vernalized (active) and vernalized (repressed) state, a decrease in interaction frequencies with sequences with low H3K27me3 and an increase in interaction frequencies with sequences with high H3K27me3 was observed. These results indicate that, once the *FLC* gene itself is repressed by polycomb, it interacts preferably with closely located polycomb targets rather than with non-polycomb targets. Furthermore, during the vernalization process a general, transient increase of interactions with sequences *in trans* was detected, which suggests changes in higher-order chromatin organization, possibly due to cold-stress induced chromatin decompaction.

Results and Discussion

Experimental approach

To elucidate early and late effects of the vernalization process on chromosomal interactions of the *FLC* gene with the rest of the genome, we performed 4C-Seq in *A. thaliana* Col-Sf2 seedlings before, during and after vernalization, in biological triplicates, using *FLC* as a bait (FIG. 1A). In seven-day-old non-vernalized seedlings (NV) the *FLC* gene is not repressed (Zhu et al., 2015). After two weeks of vernalization (2wV), non-coding RNAs are transcribed, PRC2 is assembled at the locus, and the *FLC* gene is already stably repressed in part of the cell population, while upon four weeks of vernalization (4wV) it is repressed in the majority of the cells (Angel et al., 2011). After seven additional days of growth at room temperature the level of H3K27me3 has reached its maximum and the *FLC* gene is fully repressed (FV) (De Lucia et al., 2008).

To identify sequences interacting with the *FLC* gene, a *Bgl*II fragment spanning part of the *FLC* gene body was used as a bait sequence for 4C-Seq (FIG. 1B). Sequencing was performed using Illumina technology. The analyzed *FLC* gene is located approximately at position 3.18 Mb, in an euchromatic region of the left arm of chromosome 5 (Chr5L). Euchromatin and heterochromatin have been shown to generally exclude each

other (Fransz et al., 2002; Grob et al., 2013), therefore mostly interactions with euchromatin were expected. Moreover, mapping short sequencing reads to heterochromatin is not trivial, therefore captured sequences were only mapped to unique *Bgl*II fragments within the euchromatic regions of the *A. thaliana* Col-0 genome, and pericentromeric regions were completely excluded (TAIR10, arabidopsis.org), (Bernatavichute et al., 2008). To preserve quantitative interaction frequencies and therefore allows statistical comparison of different samples for all interactions in a region of interest, interaction frequencies between the *FLC* gene and the captured sequences were calculated by normalizing the number of reads mapped to each *Bgl*II fragment to the total read coverage.

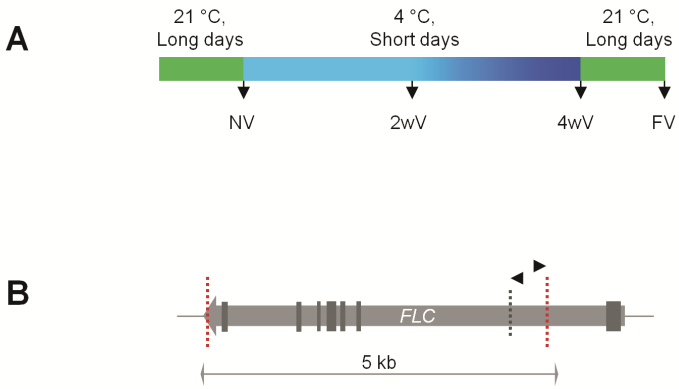


FIGURE 1. Experimental set-up for 4C. (A) Col-Sf2 seedlings grown on 1x MS-Agar plates for the indicated timespan and conditions were used for 4C analysis. (B) Graphical representation of *FLC* gene with exons indicated as dark grey boxes. A *Bgl*II fragment covering the *FLC* gene body (red dotted lines) was used as a bait for 4C. The secondary *Nla*II restriction site is marked with a grey dotted line. Primer positions for inverse PCR with 4C primers are marked with arrow heads. For details concerning the 4C technique see methods section.

Given that 4C data was obtained for each vernalization state in biological triplicates, it was possible to estimate the reproducibility of the detected interaction frequencies and to statistically quantify differences between vernalization stages. For single fragments (on average 3 kb long) or 20 kb sized windows, the obtained genome-wide interactions frequencies were too variable between biological replicates to distinguish between stage-specific conformations. The low reproducibility of 4C interaction profiles in *A thaliana* for small window sizes has been discussed before (Grob et al., 2013). Because the 4C technique allows only one ligation event with each bait sequence (van de Werken et al., 2012b) functional chromosomal interactions that are dynamic can be masked by stochastic interactions. Stochastic interactions can differ considerably, even between single cells of the same cell-type (Nagano et al., 2013). However, interactions within Chr5L displayed the reproducibility required to assess significant differences between interaction frequencies at the different vernalization stages by single-tailed, paired t-tests.

Preference for chromosomal interactions *in cis* is reduced during vernalization

The *FLC* gene is located on the left arm of chromosome 5 (Chr5L), and according to published interactome data (Grob et al., 2013), therefore an enrichment of interactions with euchromatic sequences on Chr5L was expected. Indeed, in NV seedlings, after normalization of the genome-wide 4C data to each chromosome arm length, about 40% of all sequencing reads mapped to Chr5L (FIG. 2). In case the genome-wide interaction pattern of the *FLC* gene would have been completely random, the interaction frequencies relative to the chromosome arm length would have been equally distributed over all ten chromosome arms (10% per chromosome arm). Hence, the 40% relative interaction frequency for Chr5L indicates a four-fold higher than expected enrichment. The relative interaction frequencies with the right arm of chromosome 5 (Chr5R) are similarly low as those for chromosome 1 to 4 (FIG. 2, FIG. S1). Therefore, *FLC* interactions with Chr5R are not preferred over interactions with other chromosomes, which is in agreement with chromosome arms, rather than whole chromosomes, forming territorial units in the nucleus (Berr and Schubert, 2007; Grob et al., 2013). After two weeks of cold treatment the relative interaction frequencies of *FLC* with sequences *in trans* increased uniformly, and as the reads are normalized to the total, genome-wide read coverage, the fraction of *in cis* (Chr5L) interactions decreased to about 20%. During the vernalization process, the relative interaction frequencies of *FLC* with sequences on Chr5L increased again and reached the initial level of 40% after full vernalization (FV; FIG. 2). Thus, the decrease in preference of *FLC* to interact with sequences on Chr5L is transient.

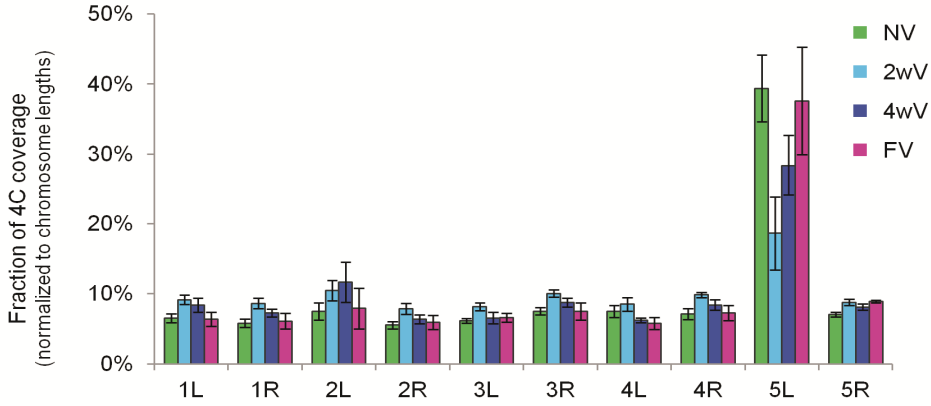


FIGURE 2. Fraction of 4C coverage for each chromosome arm. Data is normalized to genome coverage, fractionized and then normalized to the respective chromosome arm length. Hence, under complete random interaction conditions, the fraction of 4C coverage of each chromosome arm should be 10% (there are 10 chromosome arms). The majority of reads map to the left arm of chromosome 5 (5L), the chromosome arm containing *FLC*. In seedlings after 2 weeks of cold treatment the enrichment of reads on 5L is decreased, but increased again during further vernalization.

To assess if *FLC* interacts with any specific sequences *in trans*, 4C data mapped to chromosome 1 to 4 was visually screened for regions that exceeded the basal level of interaction frequency. Interestingly, all visually inspected interactions *in trans* could be tracked back to transposable elements (TEs) (data not shown). While these could indeed be specific interactions *in trans*, the exact genomic positions of these TE are ambiguous, especially because the Col-Sf2 line used in this study is not identical to the Col-0 line used for the assembled TAIR10 genome. Interestingly, it has been reported for *A. thaliana*, wheat, maize and rice that cold-stress induces elevated TE expression (Barah et al., 2013; Laudencia-Chingcuanco and Fowler, 2012; Makarevitch et al., 2015; Naito et al., 2009). The prolonged exposure of the Col-Sf2 lab strain to cold over many generations could have induced TE expression and thereby the potential promotion of TE insertions and translocations. It can therefore not be excluded that the observed, interacting TE sequences in the used Col-Sf2 line are in fact located on Chr5L. Such location would explain the elevated interaction frequencies, as the interacting TE sequences would be part of the same territorial unit. The low confidence in interactions of *FLC in trans* together with the low reproducibility of *in trans* interaction frequencies led to the decision to focus our analysis on interactions of *FLC* with sequences *in cis*.

Full vernalization induces a redistribution of preferential interactions of *FLC* within Chr5L

To visualize the general distribution of interaction frequencies between the *FLC* gene and sequences within Chr5L, the normalized 4C interaction data were used to calculate a 500 kb running mean for 100 kb sliding windows over the euchromatic part of Chr5L (FIG. 3). For all conditions studied, most intra-chromosomal interactions took place in a 500 kb window around the *FLC* gene. The intra-chromosomal interaction frequencies outside the 500 kb window were much lower and in general more similar to the average interaction frequencies observed for *in trans* interactions. The interaction frequencies within the 500 kb window around *FLC* were highest in the NV state. After two weeks of cold treatment (2wV), the frequencies were decreased ~ 2.5-fold. These lower frequencies stayed relatively low after four weeks of cold treatment (4wV), and full vernalization (FV). Analysis of the interaction frequencies outside the 500 kb window showed that in NV seedlings, these interaction frequencies were higher than the average *in trans* interaction frequencies (FIG. 3, dashed line) and decreased only slightly with increasing linear distance to the bait. After 2wV, the overall interaction frequencies outside the 500 kb window decreased to the average level of *in trans* interactions. After 4wV, the intra-chromosomal interactions outside of the 500 kb window increased again to a level comparable as observed for NV seedlings. After FV, the levels increased to an even higher level as in NV seedlings. During the cold, the general decrease in interactions on Chr5L can be explained with an overall increase of interactions *in trans* (FIG. 2). Before and after vernalization,

FLC is as likely to interact with sequences on Chr5L, however, the *FLC* interaction pattern spanning Chr5L changed between an NV and FV state (FIG. 3).

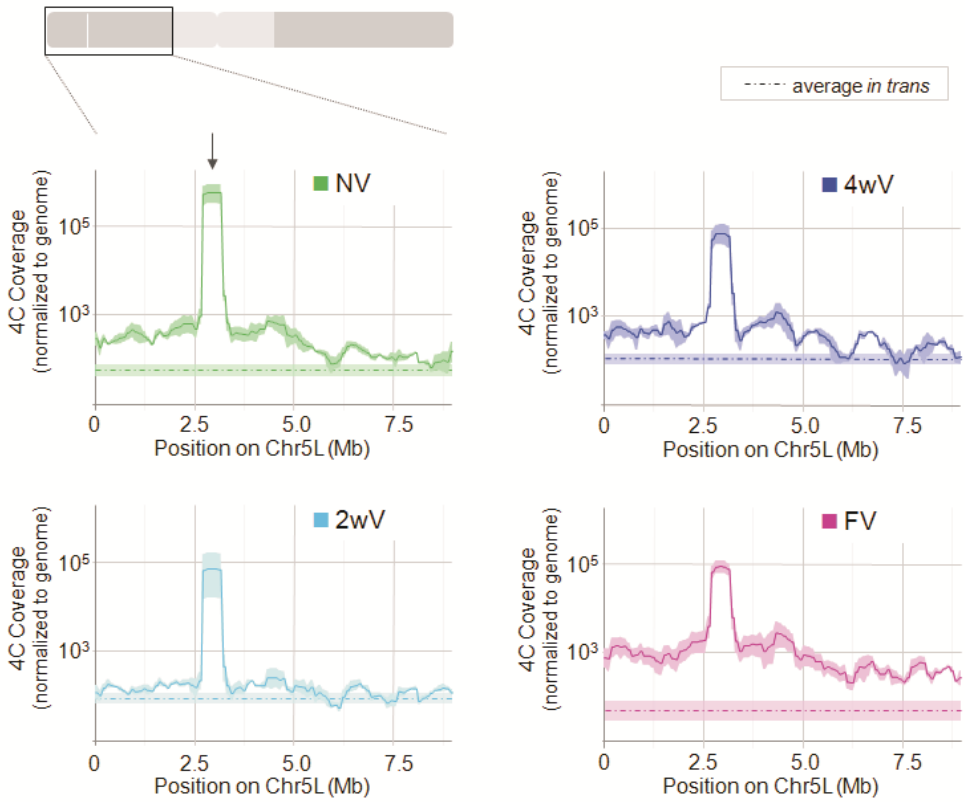


FIGURE 3. Average 4C coverage for 100 kb running mean windows on Chr5L in each vernalization state. On a schematic chromosome 5 the analyzed region (Chr5L; black box) and the *FLC* gene position (dark grey line) are indicated. Indicated are the average interaction frequencies *in cis* as a line, +/- SEM in shading. In each panel, for each vernalization state the average frequency of interactions with sequences *in trans* is shown for comparison (horizontal dashed line). For the NV graph the position of *FLC* is indicated with an arrow.

Upon vernalization frequencies of specific distant interactions increase, while frequencies of specific short-range interactions increase or decrease

To analyze the interaction patterns of *FLC* in more detail we further focused on long-range interactions between *FLC* and 20 kb regions within a 2Mb window around the *FLC* gene, and on interactions in a shorter range between *FLC* and 5 kb regions within a 200 kb window. To be able to compare interaction frequencies between the different vernalization states within Chr5L, data was normalized to the 4C coverage of Chr5L rather than the whole genome. After such normalization, the profiles of *FLC* interaction frequencies were similar for the NV and 2wV state (FIG. S2A), however, different from the profiles obtained for 4wV and FV (FIG. 4, FIG. S2B).

Using 20 kb windows, the symmetric, highly interacting domain around the *FLC* gene appeared to be approximately 400 kb wide (FIG. 4A). The width of the central domain did not change upon vernalization. In order to focus on changes in chromosomal interactions due to *FLC* silencing rather than cold treatment, for further statistical analysis only the differences between the NV and FV state were considered. In line with the earlier observation for the whole Chr5L, interaction frequencies for sequences between 200 kb and 1 Mb away from the *FLC* gene were generally higher in the FV state than the NV state. This increase was significant for several 20 kb windows distributed throughout the analyzed region (FIG. 4A).

In *A. thaliana*, regulatory chromosomal interactions are most likely to happen between sequences at a closer distance (Feng et al., 2014; Grob et al., 2014). Therefore a more detailed analysis was focused on a 200 kb window around *FLC* (FIG. 4B). Within this window, for several 5 kb regions a significant decrease in *FLC* interaction frequencies was detected between the NV and FV state. The most prominent decrease was noticed for windows directly downstream of the *FLC* gene. Additionally, regions ~40 kb downstream (+40kb) as well as ~50 kb and ~80 kb upstream (-50kb and -80kb) of the *FLC* gene interacted more frequently before than after vernalization. At the same time, for some other regions, interaction frequencies were increase from the NV to the FV state, however, only at a region ~ 55 kb downstream of *FLC* this increase was significant. Together the data indicate, that within a ~200 kb region around *FLC*, chromosomal interactions with the *FLC* gene undergo a rearrangement upon transcriptional repression of *FLC* resulting in an increased chance to interact with some sequences and a decreased chance to interact with others.

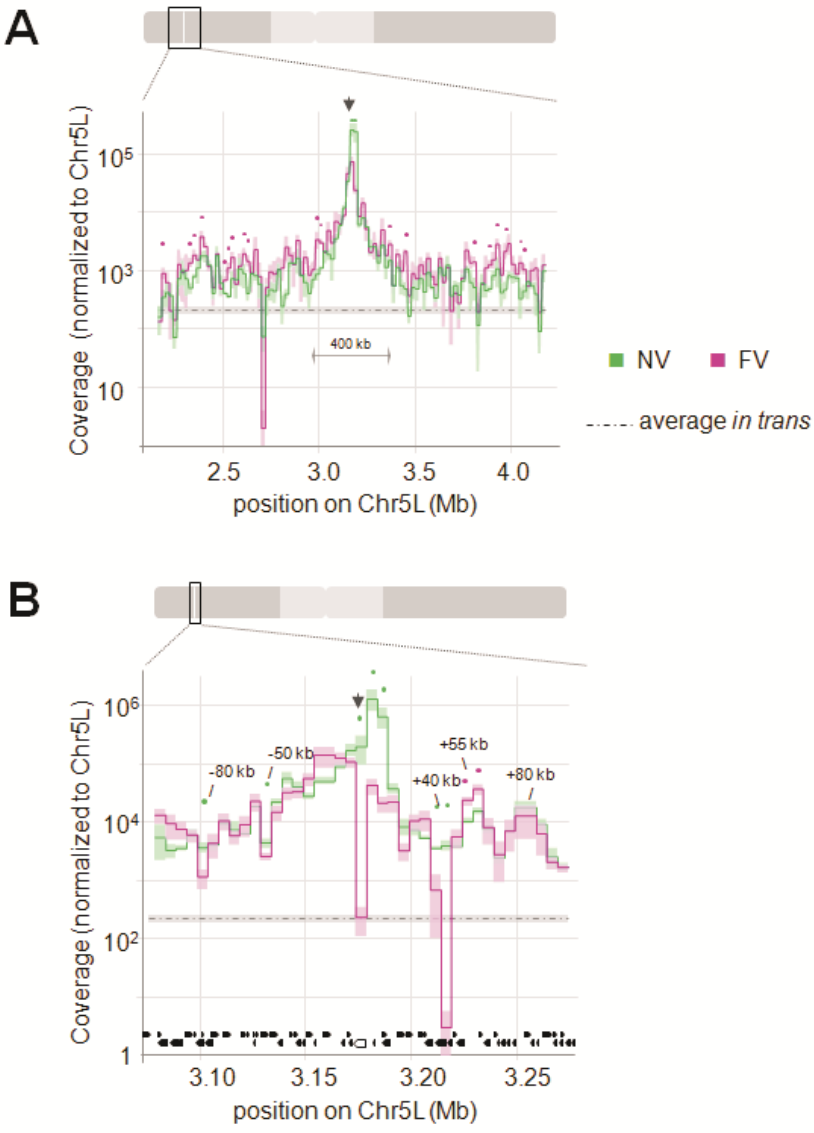


FIGURE 4. Average 4C coverage on Chr5L for distant and short-range interactions before and after vernalization. (A) Distant interactions are shown for 20 kb regions within a 2 Mb window (B) and short-range interactions are shown for 5 kb regions in a 200 kb window around *FLC*. Interaction frequencies were normalized to the total coverage on Chr5L. The position of *FLC* is pointed out with an arrow. The average interaction frequencies are indicated as a line, +/- SEM in shading. Positions with significant differences ($p<0.05$) between the interaction frequencies of the NV an FV state are marked with dots above the profiles, where the color of the dot indicates the state with the higher interaction frequency. The average +/- SEM of interactions with sequences *in trans* is indicated in grey for comparison

Only changes in short-range interaction frequencies correlate with H3K27me3 enrichment of *FLC* upon vernalization

To examine if this is true for *FLC*, publically available data sets for gene expression (AtGenExpress database) and H3K27me3 in Col-0 seedlings (Li et al., 2015) were used to investigate the transcriptional status of regions that differ or not differ in interaction frequencies with the *FLC* gene in the NV and FV stage.

For distant 20 kb regions that showed an increased interaction frequency with the *FLC* gene upon vernalization, no clear correlation was detected between the polycomb-silenced state of *FLC*, and H3K27me3 enrichment and expression levels at these regions (FIG. S3). Possibly, 20 kb windows in the *A. thaliana* genome contain too many genes to draw conclusions about the interaction with specific sequences. However, also distant single restriction fragments (on average 3 kb long) that displayed a significant increase in interaction frequency upon vernalization (FIG. S4), did not show specific enrichment for H3K27me3 marks or lowly expressed genes (data not shown). Interestingly, a 20 kb long fragment ~1.6 Mb downstream of *FLC* that showed an increased interaction frequency after vernalization contained the *KNOT ENGAGED ELEMENT 9 (KEE9)* locus (FIG. S4). KEEs, which comprise TEs of the *ATLANTYS3* and *VANDAL6* family, have been shown to interact with each other *in trans* resulting in heterochromatin islands within euchromatin (Grob et al., 2014). At this point it is unclear why the polycomb-silenced *FLC* gene would interact with such a heterochromatin domain.

For a 200 kb window around the *FLC* gene, upon transcriptional repression of *FLC*, chromosomal interactions were found to undergo rearrangements, whereby interaction frequencies increased for some sequences and decreased for others. It was hypothesized that active genes preferably interact with *FLC* before vernalization, whereas genes that preferably interact with *FLC* after vernalization are polycomb-silenced and thereby marked with H3K27me3. Indeed, the region ~55 kb downstream of the *FLC* gene that showed a significantly higher interaction frequency after vernalization overlaps with the H3K27me3-enriched region closest to *FLC* (FIG. 5). This indicates that the *FLC* gene most likely interacts with this region once both are highly enriched for the polycomb-specific H3K27me3 mark. This specific region is flanked by two genes in inverse orientation, and most likely contains the promotor region for both genes. In addition, the closest H3K27me3-enriched region ~ 70 kb upstream of the *FLC* gene displayed slightly higher interaction frequencies after vernalization. The third-closest H3K27me3-enriched region, which is located ~80kb downstream of the *FLC* gene, overlapped with a region that displayed high interaction frequencies with the *FLC* gene in both its active and repressed state (FIG. 5). We hypothesize that the genes within this region, one encoding an aminotransferase involved in embryo development and an F-box family protein that is not further annotated, are downregulated at a similar time as *FLC*. Furthermore, regions ~40 kb downstream, as well as ~50 kb and ~80 kb upstream of the *FLC* gene, with which

interaction frequencies decreased after vernalization, did not show high H3K27me3 levels (FIG. 5A). This implies that the *FLC* gene is less likely to interact with these regions once itself carries H3K27me3 and is repressed.

The increased interaction frequencies of the *FLC* gene with particular genes before vernalization might imply that these genes are highly transcribed, just as the *FLC* gene in the NV stage. However, in Col-0 these genes have only a slightly higher median level of expression than genes that did not interact with *FLC* (FIG. 5B). It has to be noted that gene expression in *A. thaliana* can vary dramatically due to e.g different growing conditions. Therefore, for a subset of genes, RNA transcript levels were measured in NV and FV Col-Sf2 seedling. Nevertheless, also in Col-Sf2 seedlings, when the expression of the *FLC* gene decreased after vernalization, transcription of the genes tested increased regardless of their position within or outside of an NV-enriched interaction region (FIG. 5C). These data suggest that the preferred contacts of the *FLC* gene with other genes before vernalization are not determined by their expression state. It has to be noted that, ideally, not transcript levels, but transcription rates of interacting genes should be measured. The measured level of RNA transcripts is determined not only by their transcription but also by their stability.

At least 20% of all coding genes in *A. thaliana* are enriched for H3K27me3 marks, and can therefore be classified as polycomb targets (Li et al., 2015; Turck et al., 2007; Zhang et al., 2007). HiC data obtained in *A. thaliana* suggests the existence of chromosomal interactions between specific H3K27me3-marked regions with elevated frequencies (Feng et al., 2014; Wang et al., 2015). In our data set, for sequences further than 200 kb away from the *FLC* gene, increased interaction frequencies after vernalization were not correlated with high H3K27me3 levels. Therefore, we did not find any indication that the repressed *FLC* gene clusters with other H3K27me3-marked polycomb regions at such longer distances. In *A. thaliana*, H3K27me3 domains are on average only 1 to 1.5 kb long and usually do not extend beyond single genes (Mozgova and Hennig, 2015). Probably this is too short to allow efficient clustering with similarly marked distant sequences. Upon vernalization, the *FLC* gene loop is disrupted (Crevillén et al., 2013). Therefore, the elevated *FLC* interaction frequencies with distant sequences after vernalization might be the result of this release of short-range interactions, rather than being sequence specific. Nevertheless, the exceedingly high frequencies for short-range interactions suggest that only at short-distance specific interactions may be stabilized. Indeed, specifically in a ~200 kb region around *FLC*, upon vernalization, interaction frequencies with H3K27me3-poor regions decreased, while interaction frequencies with H3K27me3-rich sequences increased. These results suggest that at relatively short distance, the polycomb-repressed *FLC* gene interacts more likely with polycomb than with non-polycomb targets.

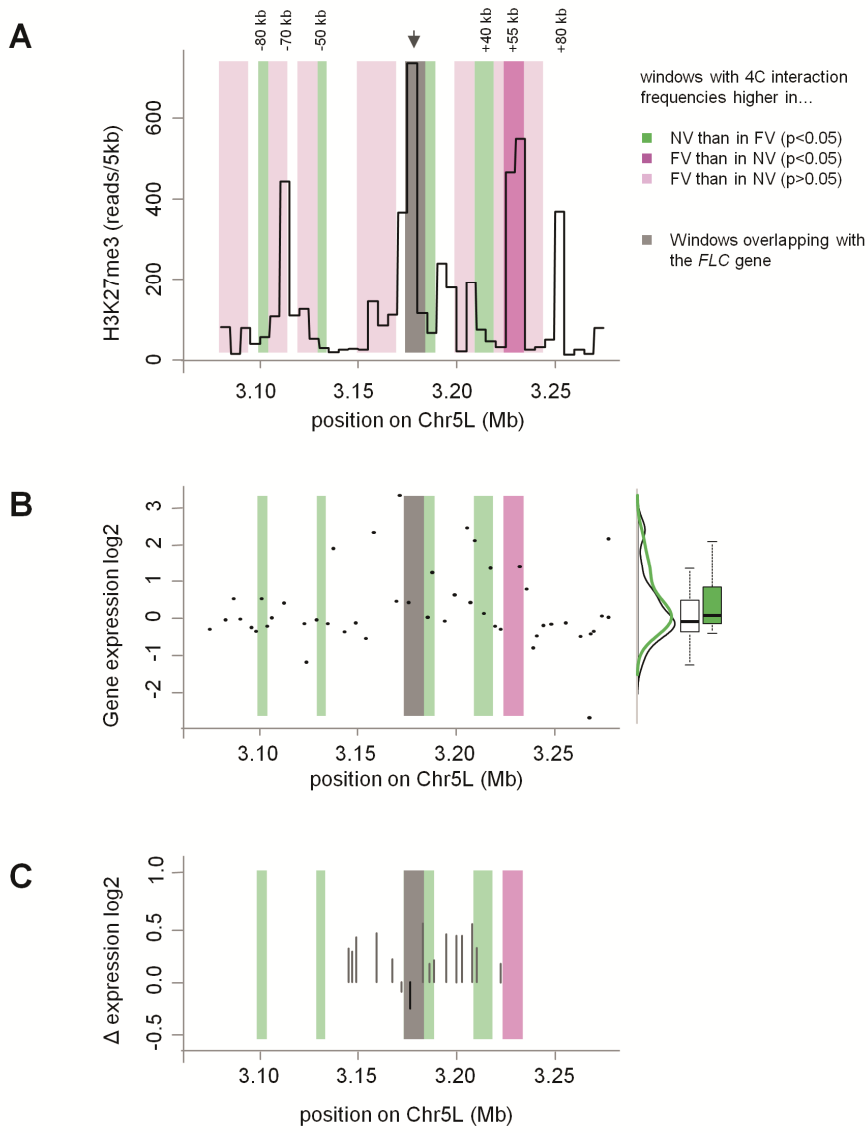


FIGURE 5. H3K27me3 levels, gene expression and changes in gene expression in a 200 kb region around *FLC* gene. The position of *FLC* is indicated with an arrow and marked with a grey window. Windows with 4C interaction frequencies significantly higher in the NV state compared to the FV state are indicated in green, windows significantly higher in the FV state in magenta, higher, but not significantly higher, in light pink. (A) Publically available H3K27me3 ChIP-seq data for 14 days old Col-0 seedlings was obtained from http://www.ncbi.nlm.nih.gov/geo. (GSM1296932; GSM1296933) (Li et al., 2015a). (B) Col-0 gene expression data was obtained from the public AtGenExpress database. The wide distribution of gene expression within windows that do have 4C interaction frequencies significantly higher in the NV than FV state and windows that do not, is displayed as a density plot as well as box plots. (C) In Col-Sf2 seedlings the fold change in gene expression from the NV to the FV state for a subset on genes measured by RT-qPCR is shown as grey bars.

Interactions of the *FLC* gene sequence with directly adjacent sequences decrease significantly upon vernalization

Within the 200 kb window around the *FLC* gene one of the most prominent changes in chromosomal interactions upon vernalization was detected. Inspection of single fragment data directly around the *FLC* gene showed that the fraction of reads mapped to the *FLC* sequence and its neighboring fragments decreased during vernalization (FIG. 6). The interaction frequency with the fragment upstream of *FLC* decreased already upon 2 weeks of cold treatment (2wV), while the frequency for the fragment downstream decreased dramatically at the end of vernalization (4wV and FV).

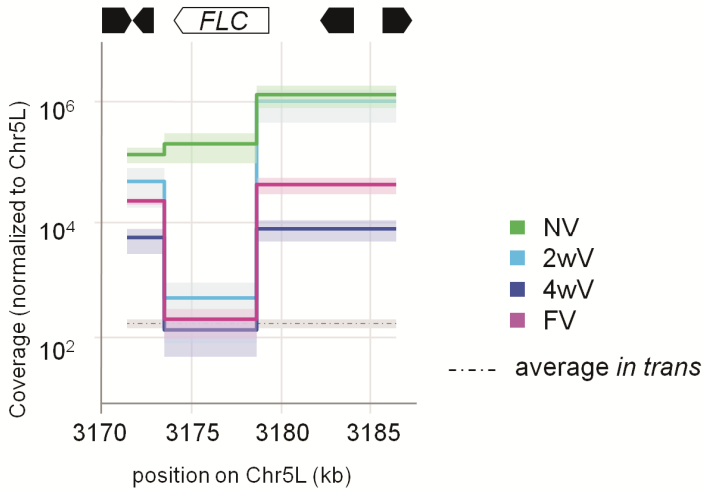


FIGURE 6. 4C interaction frequencies for individual *BglII* fragments in immediate proximity of *FLC* before, during and after vernalization. Interaction frequencies were normalized to the total coverage on Chr5L. High, differential levels of interaction frequencies at the bait itself, the fragments upstream and downstream shown as average +/- SEM. Mean +/- SEM of interactions with sequences *in trans* is indicated in grey for comparison.

Interaction frequencies measured for the bait sequence itself and adjacent sequence are indicated to be technical artifacts (van de Werken et al., 2012b). If 4C reads map to the bait itself, this is an indication that during ligation of crosslinked chromatin the bait sequence becomes circularized, rather than being ligated to available interacting sequences. 4C reads mapping to fragments up- and downstream of the bait can be picked up due to incomplete restriction digestion of crosslinked chromatin. In the study presented here, the interaction frequencies for both circularization and re-ligation did, however, change significantly between the analyzed vernalization stages, suggesting an actual difference in chromosomal interactions for the underlying sequences. In fact, a decrease in the frequency of bait circularization as well as interactions with the neighboring fragments upstream and

downstream may be explained by the disruption of the *FLC* gene loop during vernalization (Crevillén et al., 2013). When the *FLC* gene loop is present, the two *Bgl*II sites between the upstream, downstream and bait fragments would be in close proximity. After disruption of the loop when the FLC gene is in a linear conformation, the *Bgl*II sites would be more distant to each other. Therefore, the presented 4C data are in line with the existence of a vernalization-sensitive *FLC* gene loop.

The average nuclear size is elevated during cold treatment

While particular vernalization-induced changes in short-range interactions could be explained by the conformation of the *FLC* locus itself, or by a preference to cluster with similarly regulated sequences, it is unclear why overall, the fraction of chromosomal interactions of *FLC* at Chr5L decreased during cold treatment (FIG. 2). A possible cause of such a decrease in interactions *in cis* might be a change in the genome-wide compactness of the chromatin. A HiC study comparing the genome-wide interactomes of *A. thaliana* wildtype with mutants having a decreased nuclear size, showed an increase in the number of interactions *in trans*, indicating that the increase in intermingling of the chromosome territories is caused by space constraints (Grob et al., 2014). The decrease of *FLC* interaction frequencies within Chr5L during cold treatment is associated with an overall increase of interactions *in trans*, and we therefore speculated whether cold stress has an influence on chromatin organization by changing the genome-wide chromatin compactness or by changing the size of the nucleus, which in both cases might lead to intermingling of chromosome territories. To test this hypothesis, the size of nuclei of seedlings from all four stages was analyzed by flow cytometry (FIG. S5). Surprisingly, nuclei from seedlings during cold treatment (2wV and 4wV) were not smaller, but actually slightly larger than nuclei obtained from NV and FV seedlings (FIG. 7). The same trend was observed when nuclei were categorized based on their genome content into 2N up to 16N nuclei using DAPI staining (FIG. S6).

In summary, cold treatment of Col-Sf2 seedlings is associated with a slightly increased nuclear size, independent of the genome content. Hence, the increased fraction of 4C-*in trans* interactions in the cold is not caused by a decrease in nuclear size. An alternative reason might be a generally less compact chromatin state. Other types of stress, e.g. low light stress, have been shown to induce chromatin decompaction in *A. thaliana* (Tessadori et al., 2009; Florentin et al., 2013). Crosstalk has been observed between cold and light response pathways (Barah et al., 2013). We therefore hypothesize that cold stress has similar effects on chromatin condensation as low light stress and there are indeed indications this is the case (Paul Fransz, personal communication). In such a scenario, genome-wide chromatin decompaction would lead to a decondensation of all chromosome arm territories including Chr5L and hence allow more interactions of *FLC* with sequences *in trans*.

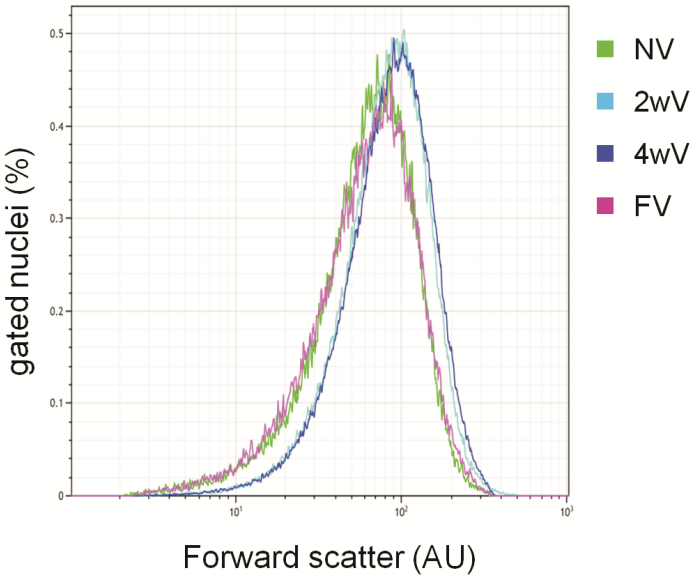


FIGURE 7. Distribution of nuclei size changes during vernalization. Nuclei size was measured as forward scatter in flow cytometry. Two independent experiments for each state were pooled resulting in more than 80 000 gated nuclei. During the cold treatment the average nuclei size is increased.

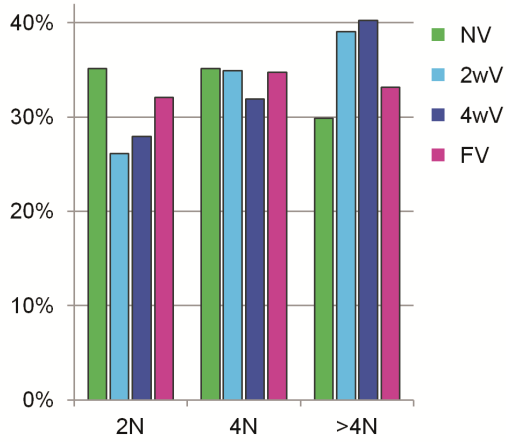


FIGURE 8. Fractions of nuclei in different ploidy levels change during vernalization. DNA content in nuclei was analyzed by flow cytometry. Two independent experiments for each state were pooled resulting in more than 80 000 gated nuclei. During cold treatment the fraction of >4N-nuclei is increasing.

The average DNA content of nuclei is elevated during cold treatment

During and after vernalization, besides the average nuclear size, also the average genome content per nucleus changed, although seedlings from all stages are morphologically very similar (FIG. S7). Non-vernalized seedlings contain ~35% nuclei with a 2N genome, ~35% nuclei with a 4N, and 30% nuclei with a > 4N genome (FIG. 8). The high fraction of polyploidy can be explained by endoreduplication, DNA replication without subsequent cell division, which is common in differentiated cells of plants with rapid life cycles (Galbraith et al., 1991; Inzé and De Veylder, 2006). During cold treatment, the fraction of diploid nuclei decreased to 26% (2wV) and 28% (4wV), while the fractions of >4N nuclei increased to 39% (2wV) and 40% (4wV) (FIG. 8). Nevertheless, after vernalization and further growth at room temperature (FV), the fraction of diploid nuclei increased again to ~32%, while the fraction of >4N nuclei decreased again to 33%. In summary, these results indicate that in *A. thaliana* Col-Sf2 seedlings, the fraction of nuclei with a higher genome copy number transiently increase during cold treatment.

The increased genome number could be due to a decrease in cell division during the cold treatment. Cold stress has been shown to inhibit e.g. microtubule synthesis, which subsequently inhibits mitosis (Chen et al., 2011; De Storme et al., 2012). After cold treatment and further growth at room temperature, the fraction of nuclei with higher genome copy number decreased again. This may be explained by the division of cells with a high genome copy number and a stalled cell cycle during the cold period. Additionally, an increase in cell proliferation, visible as seedling growth as well as the beginning of leaf organ formation, may increase the fraction of nuclei with a lower genome content.

To rule out that the temporal increase of DNA content per nucleus is due to general changes in seedling development, nuclei of seedlings grown for 10, 12 and 14 days at room temperature were subjected to the same flow cytometry analysis as seedlings subjected to a vernalization treatment. The developmental stage of seedlings grown for 10-12 days at room termperature was comparable with that of seedlings subjected to vernalization and showing a reduced growth rate during the cold treatment (FIG. S7). With seedlings grown at room temperature the fractions of nuclei with a particular genome copy number changed minimally between 10, 12 and 14 days (FIG. S8). Thus, without cold treatment no increase in genome copy number per nucleus was detected during Col-Sf2 seedling development. To test if the increase in genome copy number of nuclei is caused only by cold stress or is specific to seedlings undergoing vernalization of *FLC*, Col-0 seedlings, in which *FLC* gets rapidly repressed by the autonomous pathway independent of vernalization (Zhu et al., 2015), were subjected to a cold treatment. Genome copy numbers per nucleus were compared for Col-Sf2 and Col-0 seedlings grown for 7 and 12 days at room temperature or 7 days followed by one week of cold treatment. In both Col-Sf2 and Col-0, the fraction of 2N nuclei was lower in cold-treated seedlings compared to seedlings

that did not undergo cold treatment (FIG. S8). The exact fractions of each ploidy level in the Col-Sf2 and Col-0 seedlings was different, possibly because Col-0 showes different growth characteristics than Col-Sf2 (FIG. S7).

An increase in genome copy number per nucleus during cold treatment may affect, in addition to chromatin decondensation, the higher order organization of the chromatin. Due to the inhibition of cell division during the cold period the genome copy number per nucleus may increase and the spatial localisation of chromosomes may be affected. This is consistent with the larger fraction of *in trans* 4C interactions found in 2wV seedlings compared to NV and FV seedlings. However, throughout the cold treatment (2wV and 4wV) the fraction of polypoid nuclei remains stable, although the level of *in trans* interactions in 4wV seedlings is lower. Perhaps the decrease in interactions *in trans* towards the end of the long cold treatment might be explained by adaptation to the cold environment resulting in a slow reorganization of chromatin.

Conclusion and perspectives

The presented 4C study examined the chromosomal interaction pattern of the *FLC* gene upon its silencing induced by vernalization. The 4C interactions *in cis* appeared reproducible, which permitted the statistical analysis of the differences in interaction frequencies between the monitored stages of vernalization. Within the ~200 kb region around *FLC* gene, changes in interaction frequencies upon vernalization correlated with polycomb-repression of both the *FLC* gene as well as its interacting sequences. This observation is in line with genome-wide HiC studies that suggest the clustering of a small number of H3K27m3-enriched genomic regions. Here, we showed that the *FLC* gene interacts with such a local cluster after being silenced by polycomb. To get a better understanding what happens during the vernalization process, the genome-wide effects of the cold treatment on the chromatin organization have to be further analyzed. This can be achieved by cytological studies on chromatin compaction during cold stress as well as on endoreduplication related to the spatial distribution of chromosome territories in interphase nuclei. Ideally, H3K27me3 and gene expression levels at *FLC* interacting sequences should be measured in vernalized Col-Sf2 seedlings, because both can vary between different lines and growing conditions. Furthermore, tissue- and cell type-specific 4C analysis may allow the identification of chromosomal interactions that are specific for certain tissues or cell-types.

Material and methods

Genotypes and growing conditions

All experiments were conducted in Col-Sf2, which is a Col-0 line with a FRIGIDA introgression from Sf2 ecotype (obtained from Caroline Dean). Flow cytometry experiments were in addition performed in an in-house Col-0 line.
Bleach sterilized seeds were sown on 1xMS agar plates and stratified for 48 h at 4 °C. Seedlings were grown at 21 °C and 16 h long day conditions. For vernalization 7 day old seedlings were transferred to 4 °C and 8h short day conditions for up to 4 weeks. For full vernalization, seedlings after 4 weeks of cold treatment were transferred back to 21 °C long day conditions for 7 more days.
Of each analyzed vernalization state (NV, 2wV, 4wV, FV; FIG.1A) whole seedlings were harvested and processed as described below.

4C-Seq

3C analysis was performed similar as previously described with adaptations to *A. thaliana* (Hövel et al., 2012; Louwers et al., 2009b). 2-3 g Col-Sf2 seedling tissue was crosslinked in 2% formaldehyde in 1xPBS for 15 minutes. Nuclei were isolated, as previously described (Hövel et al., 2012; Louwers et al., 2009b), resuspended in 250 µl 1.2 x buffer B (Roche) and treated with 0.2% (w/v) SDS. SDS was inactivated by 2% (w/v) TritonX-100 and prior to restriction digestion with 400 U *Bgl*II (Roche, 40U/µl) over night. *Bgl*II was inactivated with 1.3% (w/v) SDS; SDS as inactivated with 2.8% (w/v) TritonX-100. Following, 200 µl 10x ligation buffer and 1440 µl milliQ were added (final volume 2 ml) and for intra-molecular ligation of crosslinked and digested DNA the sample was incubated with 100 U T4 ligase (Promega, 10U/µl) at 16 C for 6 hours. Subsequently, chromatin was de-crosslinked and DNA was precipitated as previously described (Hövel et al., 2012; Louwers et al., 2009b). To obtain 4C samples, DNA was digested with 50 U *Nla*III (NEB, 5U/µl) and re-ligated for circularization in a 2 ml volume as described above. 100 ng of re-ligated, precipitated DNA served as template in each inverse PCR to produce 4C libraries. PCR was conducted with the Expand Long Template PCR System (Roche) and primers that amplify sequences ligated to a *Bgl*II-*Nla*III fragment within the *FLC* gene (FIG.1B). To facilitate Illumina sequencing of 4C libraries the implemented primers also contained Illumina P5 (forward) or P7 (reverse) sequencing tails and 2-3 bp tags for multiplexing (TABLE S1). Eight individual PCR reactions for each sample were combined and purified with a PCR clean-up kit (Qiagen). Libraries for each state were mixed in equal concentration and submitted to Illumina sequencing (performed by BGI on Illumina 2000 platform). Replicates were processed in different lanes of an Illumina sequencing chip.

4C data analysis

NGS data obtained for all samples were first processed by filtering out truncated reads and the reads that did not contain the primary restriction site AGATCT (*Bgl*II). The resulting reads were mapped to a library of all AGATCT 5' and 3' fragment ends (fragments longer than 100 bp, fragment ends of 50bp length) from the TAIR10 genome used as the reference. Mapping was conducted by Bowtie2 with up to 2 mismatches allowed in the contact region only (TABLE S2). Only euchromatic regions of the Col-0 genome were analyzed further, excluding pericentromeric regions as defined in (Bernatavichute et al., 2008). Pericentromeric sequences were excluded from the analysis because of the limited confidence for exact mapping of 4C reads in heterochromatic regions. For 4C data interpretation only exact knowledge of genomic distance between a bait sequence and a captured sequence allows interpretation of physical distance. Mapping 4C data to any repetitive or multicopy sequences is ambiguous because of the limited confidence on their real genomic positions.

The mapping results were processed to obtain data concerning the total number of reads mapped to all fragments (i.e., 5' mapped reads + 3' mapped reads). These data were formatted as "wig" files for inspection in genome browser (files frag100length*.wig). Mapping results were normalized to total number of reads in each sample per 10 million reads. Data was averaged for 20 kb or 5 kb windows. Single-tailed, paired T-tests were performed on average data for 20 kb or 5 kb windows from triplicates of NV and FV state. Significantly differences between the NV and FV state were assumed for t-tests resulting in a p-value < 0.05.

In parallel, the mapping results (in *.sam files) were processed to estimat of the number of uniquely mapped reads by using eXpress (Roberts and Pachter, 2013). The resulting coverage of fragments (*.xprs files) was normalized according to the procedure based on ranking within classes of fragments (blind/non blind (where “blind” refers to captured sequences without a secondary restriction site), have short/long ends, short/long themselves) (van de Werken et al., 2012a). Ranks within each category are expressed in the scale from 0 to 1 and the total coverage of 5’ and 3’ ends is expressed as values from 0 to 2. Separately, the raw NGS data were analized by 4Cpipeline, as described by (van de Werken et al., 2012a). T-test to calculate significance between the NV and FV state were performed for single fragment data obtained from all three mapping/analysis approaches. Only fragment, for which all three data sets showed significant differences between the NV and FV state where considered for further analysis (FIG. S4).

The different normalization approaches were tested to process 4C data and to be able to compare replicates as well as different conditions. Commonly used approaches for normalization of NGS data are i) normalization to the total number of obtained reads, ii) a ranking according to coverage normalization to the region of maximum coverage iii) normalization to the region of maximum coverage (as described above). Normalization by

ranking does eliminate the exponential slope within the central peak around the bait sequence. While a ranking based normalization has the disadvantage to complicate the interpretation of the edges of the highly interactive domain around the bait, it has the advantage to possibly allow easier identification of interaction frequencies that exceed the slope. In non-ranked, exponentially decreasing interaction frequencies, such interpretation can be more difficult. Normalization to the highest peak does not impair the integrity of the exponential slope around the bait, it allows comparison of replicates but complicates comparison of states. This study focused on the comparison of interaction frequencies between the monitored vernalization stages, rather than on absolute 4C maxima. Therefore, for the majority of this study 4C data normalization to the total coverage, or the coverage for interactions *in cis* (within the chromosome arm) was used. This normalization approach preserves quantitative interaction frequencies and therefore allows statistical comparison of different samples for all interactions in a region of interest. Only to compare the fractal enrichment of 4C reads on each chromosome arm, the sum of all interaction frequencies for each chromosome arm was normalized to the particular chromosome arm length.

Processing public data sets

Gene expression data for all genes on chromosome 5 within a 2 Mbp range around *FLC* was obtained from the AtGenExpress database for expression in seedlings by using the E-Northern Webtool (http://bar.utoronto.ca/affydb/cgi-bin/affy_db_exprss_browser_in.cgi, Toufighi et al., 2005). Gene expression data was originally obtained from young (stage 1.0) and older (stage 1.9) Col-0 seedlings, grown on MS plate.

H3K27me3 ChIP-seq data for 14 days old Col-0 seedlings (GSM1296932; GSM1296933), mapped to TAIR 10 genome, was obtained from http://www.ncbi.nlm.nih.gov/geo, (Li et al., 2015a). For further analysis H3K27me3 ChIP-Seq read counts from a biological duplicate were pooled and subsequently summed up for 5 kb-bins along chromosome 5.

RT-qPRC

RNA was isolated with TRIzol® Reagent and a DNase Kit following the manufacturer's instructions (ThermoFisher) from NV and FV seedlings. cDNA was synthesized from RNA with the RT minus Kit (ThermoFisher). Following cDNA was analyzed by quantitative PCR using HOT FIREPol EvaGreen qPCR Mix Plus (Solis BioDyne #08-24-00001) and primers listed in TABLE S3.

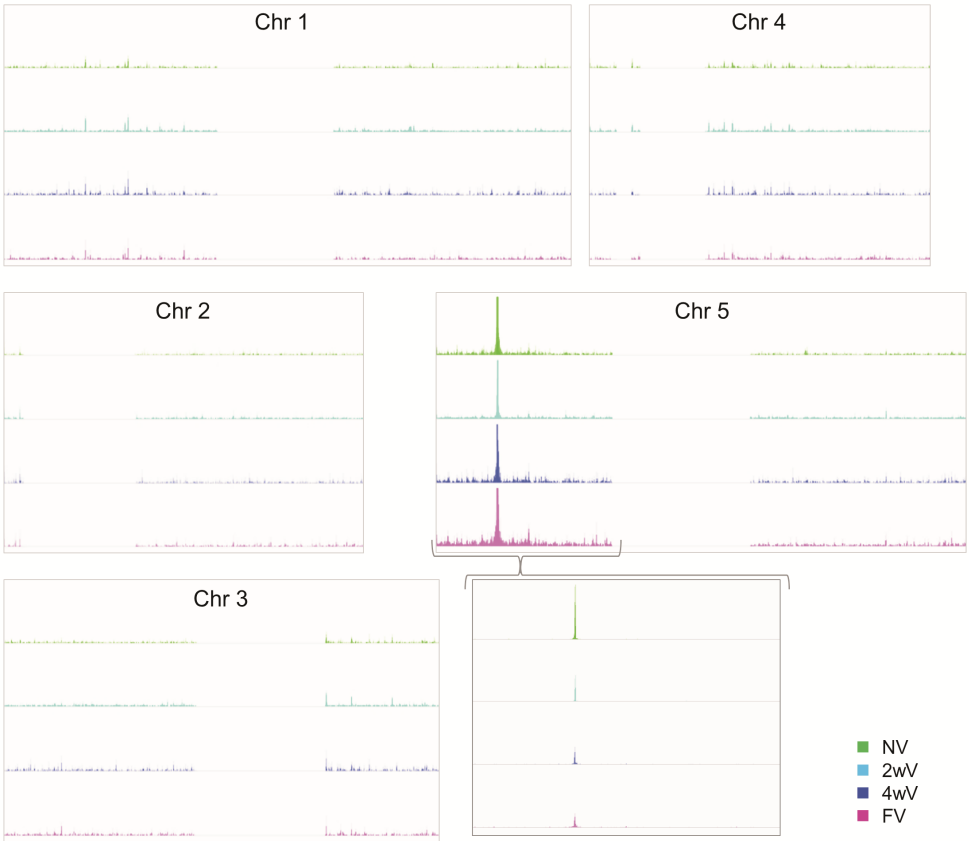
Flow cytometry

Seedlings were harvested and washed twice in water to remove agar residues. From washed and decanted seedlings, nuclei were isolated as described previously (Galbraith et al., 1983), stained with DAPI and analyzed by flow cytometry.

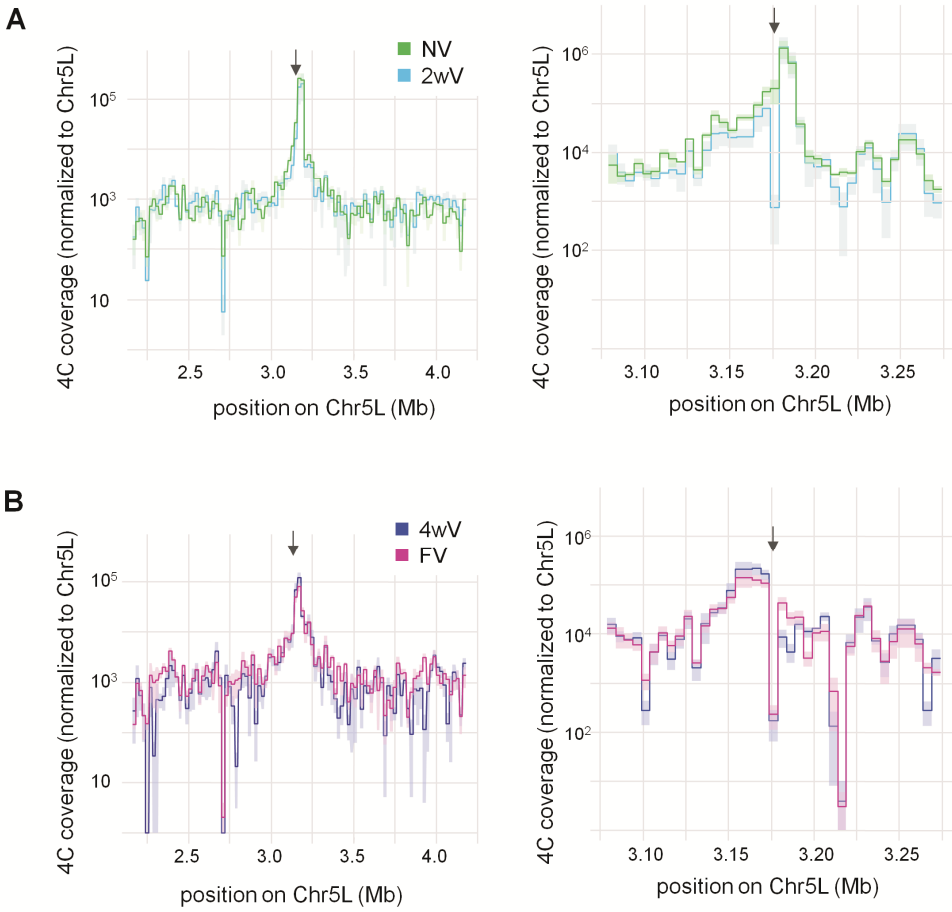
Acknowledgements

We would like to express our gratitude to Caroline Dean, as well as current and former members of her lab, for providing seeds, as well as extensive consultation on the matter of *FLC* and vernalization. Furthermore, we would like to thank Willem Stiekema and Paul Fransz for comments on the manuscript.

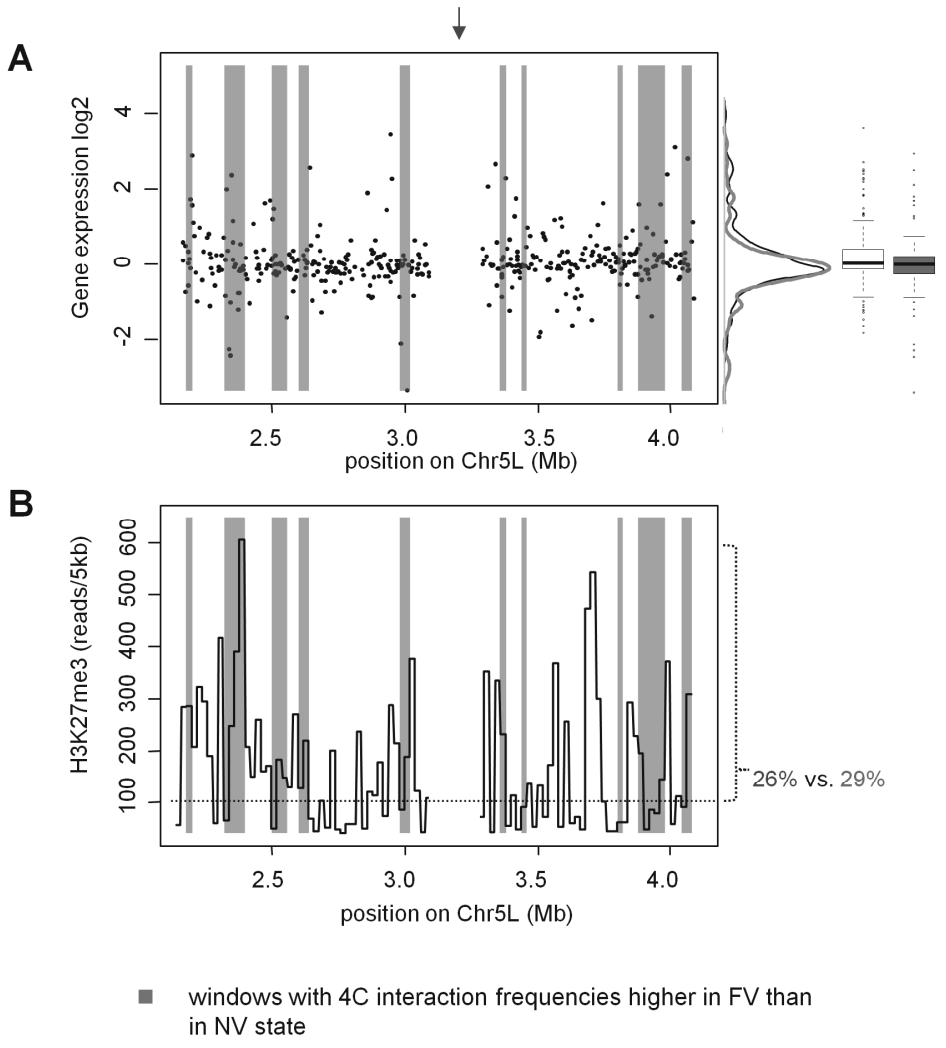
Supporting Information



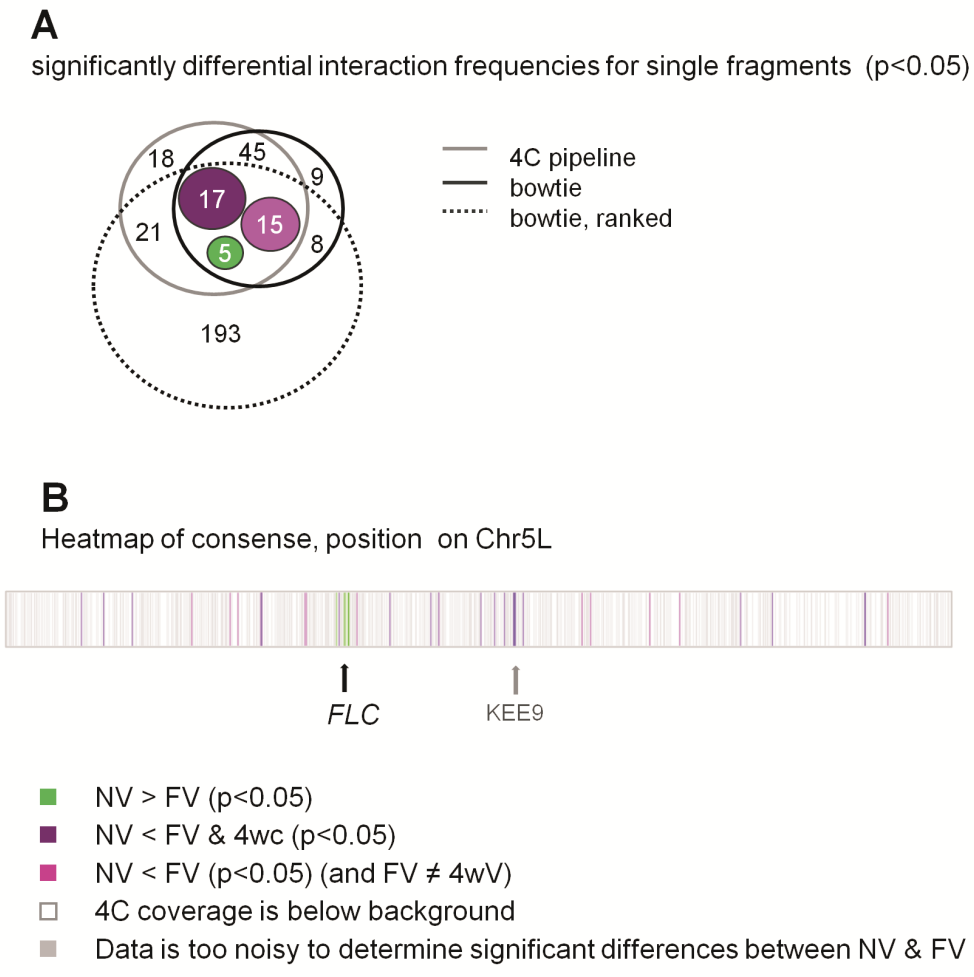
SUPPLEMENTAL FIGURE 1. 4C *FLC* interaction frequencies of 20 kb windows on all 5 chromosomes excluding pericentromeric regions. 4C data was normalized to genome-wide coverage and average +/- SEM for biological triplicates was calculated. The range of the y-axis is 0 to 20000 reads, cutting the maximum coverage on Chr5L off, to allow showing the overall genomic level of coverage. Below Chr5 a zoom into Chr5L with a maximum of 400000 reads/window is shown to underline the preference for short-range interactions.



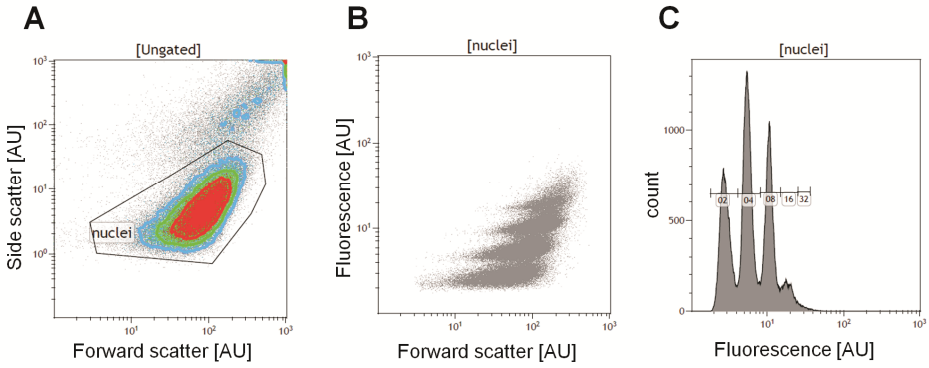
SUPPLEMENTAL FIGURE 2. Comparison of 4C coverage on Chr5L between NV and 2wV or 4wV and FV, respectively. Interaction frequencies were normalized to total coverage on Chr5L. The position of *FLC* is indicated with an arrow. (A) Window size 2 Mb around *FLC*, showing an overlay of 20 kb window means for NV & 2wV, and 4wV & FV; +/- SEM. (B) Window size 200 kb around *FLC*, showing an overlay of 5 kb window means for NV & 2wV, and 4wV & FV; +/- SEM.



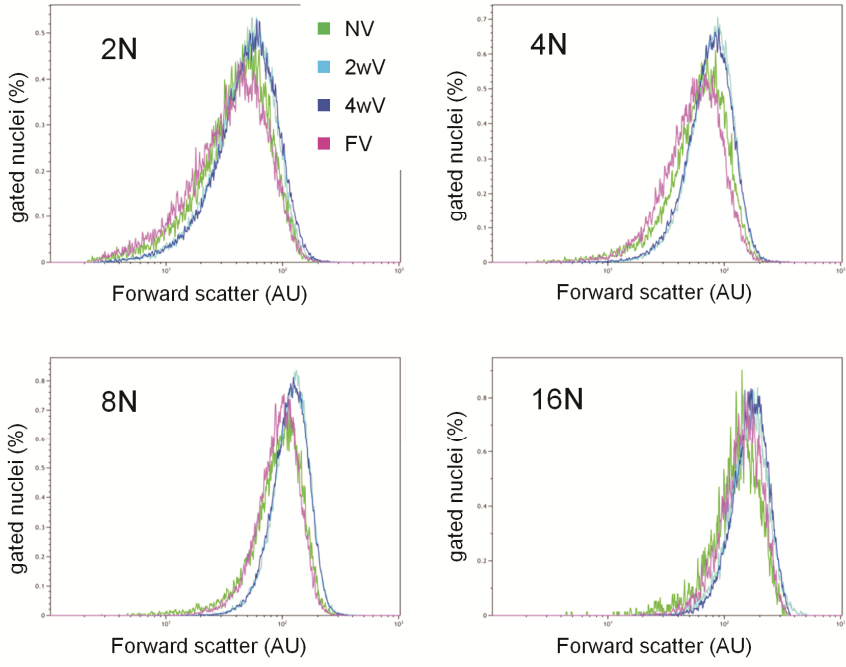
SUPPLEMENTAL FIGURE 3. Gene expression (A) and H3K27me3 levels (B) in 2 Mb region around *FLC* gene. The position of *FLC* is indicated with an arrow. 20 kb regions with 4C interaction frequencies significantly higher in FV state compared to NV state are indicated in grey. (A) Gene expression data was obtained from the public AtGenExpress database. The wide distribution of gene expression levels within windows that do have 4C interaction frequencies significantly higher in FV compared to NV and windows that do not, is displayed in a density plot as well as box plots. (B) Publically available H3K27me3 ChIP-seq data for 14 days old Col-0 seedlings was obtained from http://www.ncbi.nlm.nih.gov/geo. (GSM1296932; GSM1296933) (Li et al., 2015a). The fraction of H3K27me3 enriched regions (above background, dotted line) within windows that do have 4C interaction frequencies significantly higher in FV compared to NV is 29%, while the fraction of H2K27me3 enriched regions that do not overlap with these windows is 26%. Regions in a 200 kb window around *FLC* were omitted from this analysis on distant sequences. For analysis of short range seqeuence, see FIG. 5.



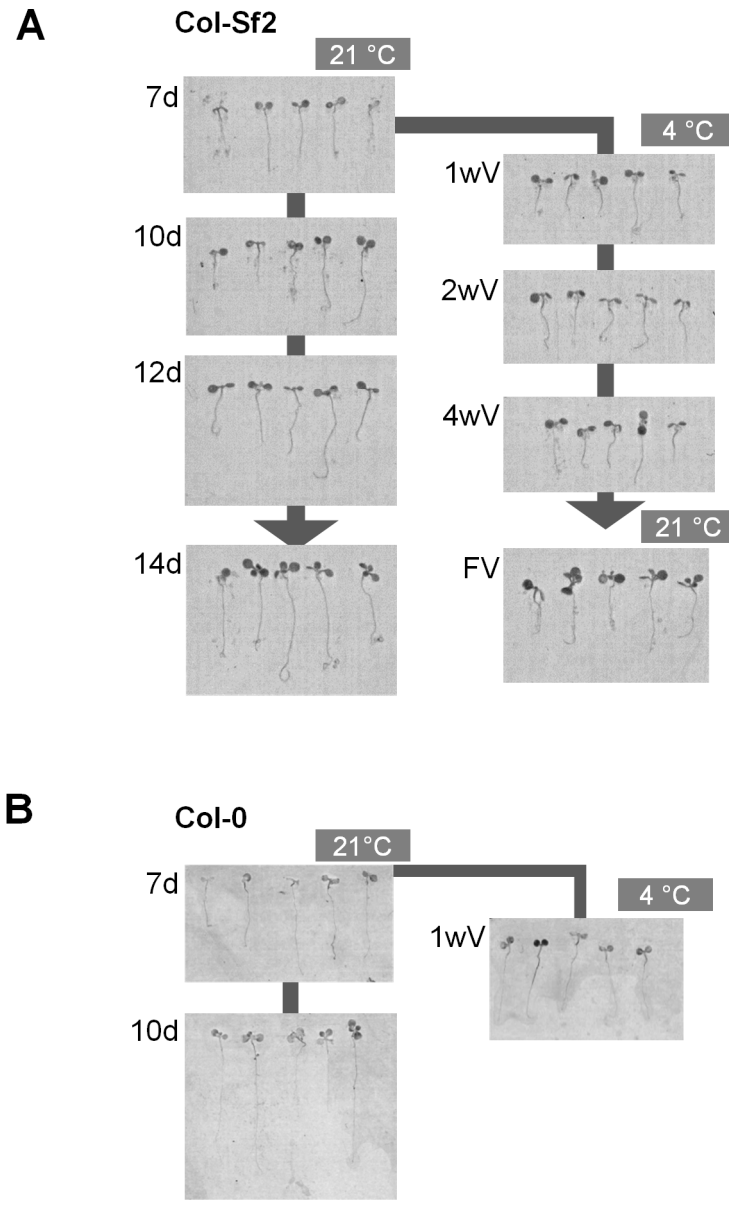
SUPPLEMENTAL FIGURE 4. Comparing significant differences in interaction frequencies between NV and FV state for single *Bgl*II fragments on Chr5L measured by different mapping/4C analysis approaches. (A) Venn diagram indicating how many fragments on Chr5L showed differential 4C coverage (p<0.05) and what fragments are identified independently of the 4C mapping approach (as indicated in legend). (B) Positions of single fragments of the differential categories indicated in (A) on Chr5L. Positions with higher interaction frequencies before vernalization cluster around the viewpoint, similar as it has been seen for standardized 5 kb windows (Fig 4b). Positions with higher interaction frequencies after vernalization are spread out over the chromosome arm.



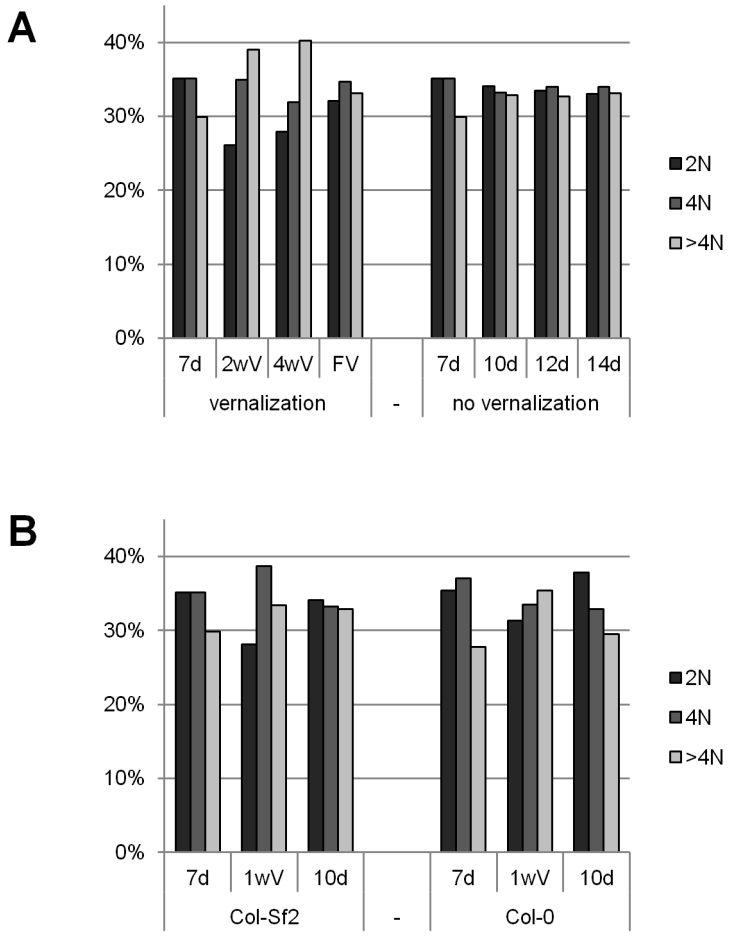
SUPPLEMENTAL FIGURE 5. Example of flow cytometry set-up to measure DNA content in single nuclei. (A) Nuclei were gated by the correlation of forward ad side scatter. (B) Nuclei size was estimated by forward scatter and DNA content was measured by DAPI staining. (C) Nuclei were categorized by their DNA content. Two independent experiments for each state were pooled resulting in more than 80 000 gated nuclei.



SUPPLEMENTAL FIGURE 6. Distribution of nuclei size in different ploidy classes changes during vernalization. Nuclei size was estimated by forward scatter in flow cytometry and DNA content was measured by DAPI staining. Two independent experiments for each state were pooled resulting in more than 80 000 gated nuclei. During the cold treatment the average nuclei size is increased in 2N, 4N, 8N and 16N nuclei.



SUPPLEMENTAL FIGURE 7. Seedling morphology. Col-Sf2 seedlings were either grown in normal warm conditions for 7 to 14 days, or after 7 days of normal conditions subjected to vernalization treatment. Col-0 seedlings were grown for 7 to 10 days in normal warm conditions or for 7 days and subsequently for one week at cold (veranlization) conditions. For each stage five seedlings were randomly picked from MS plates and photographed.



SUPPLEMENTAL FIGURE 8. Fractions of nuclei in different ploidy levels in cold treated and not cold treated seedlings. Ploidy level was analyzed once by flow cytometry for nuclei from Col-Sf2 grown in normal warm conditions (A) as well as from Col-Sf2 and Col-0 briefly exposed to cold (B). The resulting nuclei fractions with different DNA content levels are compared to the data obtained for Col-Sf2 seedlings during vernalization, as shown in Fig 9.

TABLE S1. 4C primers

		Forward primer * (P5 in capitals), tag in grey, *Bgl*II site in italics
Sequencing lane 1	Index 1: NV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTaccccatagcaactctat*agatct*
	Index 2: 2wV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTgtaccccatagcaactctat*agatct*
	Index 3: 4wV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTcaaaccccatagcaactctat*agatct*
	Index4: FV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTgataccccatagcaactctat*agatct*
Sequencing lane 2	Index 1: NV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTgataccccatagcaactctat*agatct*
	Index 2: 2wV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTcaaaccccatagcaactctat*agatct*
	Index 3: 4wV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTgtaccccatagcaactctat*agatct*
	Index4: FV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTaccccatagcaactctat*agatct*
Sequencing lane 3	Index 1: NV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTcaaaccccatagcaactctat*agatct*
	Index 2: 2wV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTgataccccatagcaactctat*agatct*
	Index 3: 4wV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTaccccatagcaactctat*agatct*
	Index4: FV	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGC TCTTCCGATCTgtaccccatagcaactctat*agatct*

* indicated forward primers are all used in combination with the same reverse primer CAAGCAGAAGACGGCATACGAtgtcttttgtcacacaaccttg (P7 in capitals)

TABLE S2. Results of NGS data processing

sample	data file	number of reads			% reads mapped to eu-chromatin
		total	after filtering	mapped to fragment library	
Index 1: NV	arabidopsis_1_1.fq	16871432	16126982	9487926	58.83%
Index 2: 2wV	arabidopsis_2_1.fq	14878396	14448008	9029675	62.50%
Index 3: 4wV	arabidopsis_3_1.fq	10365781	10064278	7737413	76.88%
Index4: FV	arabidopsis_1_2.fq	11967845	11501816	7133080	62.02%
Index 1: NV	arabidopsis_2_2.fq	16700212	15832576	9873466	62.36%
Index 2: 2wV	arabidopsis_3_2.fq	11375020	10934101	6597118	60.34%
Index 3: 4wV	arabidopsis_1_3.fq	5694192	5195956	3526168	67.86%
Index4: FV	arabidopsis_2_3.fq	8961788	8648700	4957468	57.32%
Index 1: NV	arabidopsis_3_3.fq	11459496	11241970	6795827	60.45%
Index 2: 2wV	arabidopsis_1_4.fq	12623982	11859951	5446140	45.92%
Index 3: 4wV	arabidopsis_2_4.fq	11210240	10990650	7591159	69.07%
Index4: FV	arabidopsis_3_4.fq	10702878	10516643	6514129	61.94%

TABLE S3. Primers implemented in RT-qPCR

Gene	Forward primer	Reverse primer
At5g10050	CCGCCATGGAAAACACATTCA	TGGACCAGGTGCCATAACAG
At5g10060	AGATGTTCCTCTGCCACTTGA	TGCTGATGCTATCTTCTCAGCC
At5g10070	GGACTGATAATCTTGCAAGGGAAG	ACCTGTGCACCTCTTTGCA
At5g10100	TCCACACAGCAACAACTCAAC	ATGGGAGAGAGAGTACCGTCA
At5g10110	ACAGGAGATGACTTGCTTAGATCA	AGTTGCTAGCTGCTTTCTGACT
At5g10130	ATGCGTTTTGGTTCTCCCAC	TCTTGCCCCGCGGATATATTG
At5g10140	AAACGTCGCAACGGTCTCAT	ATCTTGACCAGGTTATCGCCG
At5g10150	TGTGATGAACCGGCTTACGAT	ATTCAGCACAGTCAGAAGGGT
At5g10160	TCCCTATCTCCACCGTTGCT	CGGAAACGCCTCATACCTTAG
At5g10170	ACTGTGAAGCCAAAGACTGTGA	GCCCAAGATATTCCCTCACG
At5g10180	TCTGCATTCCGCAGAGCATT	CCGGTCCGATTGCTATCTCT
At5g10190	CTCTCGGTGCTTTCCTCTGG	GACTAGGGACTGAATCGCGG
At5g10200	CGCAGAGAGGTTGGTGAAGA	CTCTCACTCGCTGTTTCCCA
At5g10220	TTCAGTGGTCGGAGAATGGC	CGTTGCGTTTCTATGAGCCAAA
At5g10230	CTTTGAGCGAGCTGTGATGTTG	GGCATTAAAGAGTTCAAGAGCAG
actin	GGTAACATTGTGCTCAGTGGTGG	AACGACCTTAATCTTCATGCTGC