Histone H1 protects telomeric repeats from H3K27me3 invasion in *Arabidopsis*

Gianluca Teano[1], Lorenzo Concia[1], Léopold Carron[2], Léa Wolff[1], Kateřina Adamusová[3,4], Miloslava Fojtová[3,4], Michael Bourge[5], Amira Kramdi[1], Vincent Colot[1], Ueli Grossniklaus[6], Chris Bowler[1], Célia Baroux[6], Alessandra Carbone[2], Aline V. Probst[7], Petra Procházková Schrumpfová[3,4], Jiří Fajkus[3,4], Simon Amiard[7], Stefan Grob[6], Clara Bourbousse[1,*] and Fredy Barneche[1,*]

[1]Institut de biologie de l'Ecole normale supérieure (IBENS), Ecole normale supérieure, CNRS, INSERM, Université PSL, Paris, France

[2] Sorbonne Université, CNRS, IBPS, UMR 7238, Laboratoire de Biologie Computationnelle et Quantitative (LCQB), 75005, Paris, France

[3]Mendel Centre for Plant Genomics and Proteomics, Central European Institute of Technology, Masaryk University, Brno, Czech Republic

[4]Laboratory of Functional Genomics and Proteomics, NCBR, Faculty of Science, Masaryk University, Brno, Czech Republic

[5]Imagerie-Gif, Plateforme de Cytométrie, CNRS - I2BC, Gif sur Yvette, France

[6]Department of Plant and Microbial Biology & Zürich-Basel Plant Science Center, University of Zürich, Switzerland

[7]CNRS UMR6293, Université Clermont Auvergne, INSERM U1103, GReD, CRBC, Clermont-Ferrand, France

* Correspondance: clara.bourbousse@bio.ens.psl.eu; barneche@bio.ens.psl.eu

Abstract

Linker histones play a pivotal role in shaping chromatin architecture, notably through their globular H1 (GH1) domain that contacts the nucleosome and linker DNA. Yet, the interplay of H1 with chromatin factors along the epigenome landscape is poorly understood. Here, we report that *Arabidopsis* H1 favors chromatin compaction and H3K27me3 marking on a majority of *Polycomb*-targeted protein-coding genes while preventing H3K27me3 accumulation on telomeres and pericentromeric interstitial telomeric repeats (ITRs). These contrasting effects of H1 on H3K27me3 enrichment are associated with long-distance effects on the 3D organization of telomeres and ITRs. Mechanistically, H1 prevents ITRs from being invaded by Telomere Repeat Binding 1 (TRB1), a GH1-containing telomere component with an extra-telomeric function in targeting *Polycomb* to genes bearing telomeric motifs. We propose that reciprocal DNA binding of H1 and TRB1 to clustered *telobox* motifs prevents H3K27me3 accumulation on large chromosomal blocks, conferring a sequence-specific role to H1 in epigenome homeostasis.

Running title: H1 and H3K27me3 interplay on telomeric repeats

Introduction

Besides core histones, chromatin architecture and functionality rely on linker histone H1 whose central globular (GH1) domain sits on the nucleosome dyad while its intrinsically disordered carboxy-terminal domain binds linker DNA at the nucleosome entry and exit sites (Bednar et al., 2017; Zhou et al., 2015). H1 incorporation directly influences the physico-chemical properties of the chromatin fiber and further modulates nucleosome distribution and chromatin compaction. Therefore, H1 contributes to the local variation in transcriptional activity by affecting the accessibility of transcription factors and RNA polymerases to chromatin but also through interactions with histone and DNA modifiers (reviewed in Bednar et al., 2016; Fyodorov et al., 2017; Hergeth and Schneider, 2015).

Polycomb-Group (PcG) proteins are other important determinants of chromatin compaction and transcriptional activity, influencing cell identity and differentiation in metazoans (Grossniklaus and Paro, 2014; Schuettengruber et al., 2017), plants (Hugues et al., 2020) and unicellular organisms (Schubert, 2019). In metazoans, the chromatin of PcG target genes is highly compacted (Francis et al., 2004; Shao et al., 1999; Shu et al., 2012), a feature thought to hinder transcription (reviewed in Illingworth, 2019; Schuettengruber et al., 2017). The repressive activity of PcG proteins on transcription involves the enzymatic activity of *Polycomb* Repressive Complex 1 (PRC1) and 2 (PRC2) mediating histone H2A Lysine monoubiquitination (H2Aub) and histone H3 Lysine 27 trimethylation (H3K27me3), respectively (Grossniklaus and Paro, 2014; Schuettengruber et al., 2017).

Both nucleosomal and higher-order chromatin organization rely to a large extent on the regulation of chromatin compaction and accessibility (Bonev and Cavalli, 2016; Santos et al., 2020) in which both H1 and PRC2 play essential roles (Feng et al., 2014; Geeven et al., 2015; Grob et al., 2014; Liu et al., 2016; Moissiard et al., 2012; Sexton et al., 2012; Sun et al., 2020). Firstly, PRC1 subunits such as Posterior sex combs (Psc) in Drosophila, Chromobox 2 (Cbx2) in mammals, or EMBRYONIC FLOWER1 (EMF1) in *Arabidopsis* display highly positively charged regions that trigger chromatin compaction *in vitro* (Beh et al., 2012; Grau et al., 2011) and can mediate gene silencing and affect genome topology *in vivo* (Lau et al., 2017; Terranova et al., 2008). Secondly, PRC2 favors chromatin compaction, either by promoting PRC1 recruitment or through its subunit Enhancer of Zeste homolog 1 (Ezh1) in a mechanism not necessarily relying on the H3K27me3 mark (Margueron et al., 2008). Specific functional interplays between PcG and H1 in chromatin compaction have also begun to emerge (Yuan et al., 2012). In human, the preferential interaction of H1.2 with H3K27me3 nucleosomes promotes chromatin compaction (Kim et al., 2015), while, conversely, PRC2 displays substrate preferences for H1-enriched chromatin fragments *in vitro* (Martin et al., 2006).

In *Arabidopsis thaliana*, two canonical H1 variants, H1.1 and H1.2 hereafter referred to as H1, are ubiquitously expressed in somatic cells and display similar chromatin association properties while an atypical H1.3 variant is only expressed under stress conditions or in specific cellular contexts (reviewed in Kotliński et al., 2017; Over and Michaels, 2014; Probst et al., 2020). H1 is enriched over transposable elements (TEs), which display high nucleosome occupancy and various heterochromatic hallmarks, such as H3K9me2 and heavily methylated cytosines (Choi et al., 2019; Rutowicz et al., 2019; Wollmann et al., 2017). In contrast, H1 is less abundant over genes marked by transcriptionally permissive histone modifications (Rutowicz et al., 2015; Choi et al., 2019). Hence, as in other eukaryotes, incorporation of *Arabidopsis* H1 is thought to dampen transcription elongation, an effect that also applies to the production of TE-derived short interfering RNAs (siRNAs) (Papareddy et al., 2020), thereby not only restricting RNA Polymerase II (Pol II) but also RNA Pol IV activity. *Arabidopsis* H1 further restricts accessibility of DNA methyltransferases and demethylases that target TE sequences to regulate their heterochromatinization and silencing (He et al., 2019; Liu et al., 2020; Lyons and Zilberman, 2017; Wollmann et al., 2017; Zemach et al., 2013). Interestingly, DNA methylation has been found to be prevalently anti-correlated with H3K27me3 along the epigenome of *Arabidopsis* somatic cells (Zhang et al., 2007), but there is currently little information on the role played by H1 in this antagonism and, more generally, in the dynamic equilibria modulating chromatin compaction and accessibility in plants.

Using cytogenetic analyses we recently reported that *Arabidopsis H1* knockout plants present globally altered H3K27me3 levels whilst, intriguingly, a few discrete subnuclear foci of undetermined nature oppositely display increased H3K27me3 signals (Rutowicz et al., 2019). Changes in the H3K27me3 chromatin landscape may underlie the effect of developmentally programmed H1 depletion during specific cellular transitions. Indeed, massive degradation of H1 during the formation of *Arabidopsis* mega- and microspore mother cells coincides with heterochromatin loosening but also with a reduction in H3K27me3 nuclear signals (She et al., 2013; She and Baroux, 2015; He et al., 2019; Hsieh et al., 2016). In this study, we set out to determine how *Arabidopsis* H1 impacts the genomic distribution and accessibility of H3K27me3. We report that H1 globally favors H3K27me3 enrichment and low accessibility of PRC2 target genes and, conversely, prevents the accumulation of this mark over telomeres and pericentromeric interstitial telomeric repeats while also influencing their 3D organization in the nuclear space. We explored the specificity of telomere-associated chromatin alterations through the possible involvement of Telomeric Repeat Binding 1 (TRB1), a GH1-containing Myb-histone protein abundant on telomeres (Schrumpfová et al., 2014) that also mediates H3K27me3 deposition on genes with telomeric motifs acting as *Polycomb* response elements (Schrumpfová et al., 2016; Xiao et al., 2017; Zhou et al., 2018). Collectively, these findings lead us to propose that H1 contributes to the control of H3K27me3 homeostasis between structurally distinct chromosomal regions by modulating PRC2 activity in a sequence-specific manner through competitive DNA binding with TRB proteins.

Results

H1 is abundant in the body of H3K27me3-marked genes and reduces their accessibility

To explore the relationship of H1 and H3K27me3 and its links to the chromatin accessibility landscape, we first compared the genomic distribution of H1 with those of H3K27me3. To maximize H1 ChIP-seq specificity, we used a GFP-tagged version of the most abundant variant *H1.2* transcribed under the control of its own promoter in wild-type (WT) plants (Rutowicz et al., 2015). In agreement with previous studies in plants and other eukaryotes (Cao et al., 2013; Choi et al., 2019; Izzo et al., 2013; Rutowicz et al., 2015), H1 distribution covered most of the *Arabidopsis* genome without displaying clear peaks. However, examination of H1 profiles over protein-coding genes, which are the main carriers of H3K27me3 in the *Arabidopsis* genome (Roudier et al., 2011; Sequeira-Mendes et al., 2014; Wang et al., 2014), showed that H1 is highly abundant on H3K27me3-marked gene bodies, especially towards their 5' region (Figure 1A). As concluded from the analysis of H3 profiles, such a strong effect was not due to increased nucleosome occupancy (Figure S1A).

Specific examination of transcription start sites (TSS) confirmed that H1 is more abundant on H3K27me3-marked genes than on genes displaying histone hallmarks of active Pol II initiation (H3K4me3), elongation (H2Bub) or having high transcript levels (Figure S1A). H1 also appears to be globally much more abundant on H3K27me3-marked genes than on heterochromatic transposable elements (TEs) (Figure 1A), which are themselves H1-rich and heavily condensed (Choi et al., 2019). These genome-wide correlations suggest the existence of functional interplays between H1 and PRC2 activity on protein-coding genes.

We then examined the chromatin accessibility of H3K27me3-marked genes by Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) analysis of 4C nuclei, the most abundant ploidy level in our sample preparation. In WT plants, H3K27me3-marked genes displayed very low chromatin accessibility as compared to expressed genes (Figure 1B), which typically produce a sharp peak at their TSS in such analyses (Lu et al., 2016). We similarly examined *h1.1 h1.2* (*2h1*) double mutant plants to test whether H1 contributes to this low accessibility. Oppositely to transcriptionally active genes, H3K27me3-marked genes display no change of the TSS accessibility peak but they are globally more accessible along their body and over their promoter and terminator domains (Figure 1B and S1B). Hence, H1 appears to be functionally related to H3K27me3 marking in term of local abundance and chromatin accessibility.

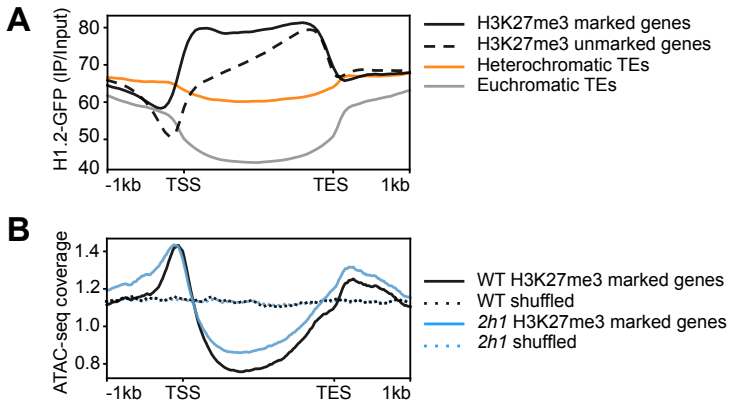


Figure 1. H1.2-GFP is enriched over PRC2-targeted chromatin and contributes to restrain its accessibility. A. H1.2-GFP median level over different genetic contexts. H3K27me3-marked genes defined in our ChIP-seq analysis (n=7542) are compared to all other annotated protein-coding genes. Heterochromatic versus euchromatic TEs were defined according to Bernatavichute et al., (2008). **B.** ATAC-seq analysis of chromatin accessibility of H3K27me3-marked genes in WT and *2h1* nuclei with similar ploidy levels (4C). Mean read coverage (reads per million) is used as a proxy of chromatin accessibility. Dashed lines represent ATAC-seq read density over random permutations (shuffled) of the genomic locations of genes marked by H3K27me3 in WT and *2h1* nuclei. TSS, transcription start site. TES, transcription elongation stop.

H1 is necessary for H3K27me3 enrichment and low chromatin accessibility at a majority of PRC2 targeted genes

To assess the mechanistic links between H1 and H3K27me3, we first profiled the H3K27me3 landscape upon H1 depletion. Considering the reported ~2-fold decreased abundance of H3K27me3 level in *2h1* seedlings (Rutowicz et al., 2019), we employed a spike-in ChIP-seq approach (ChIP-Rx) using a fixed amount of Drosophila chromatin exogenously added in each sample prior to immunoprecipitation and subsequently quantified in input and immunoprecipitated DNA samples (as in Nassrallah et al., 2018). In line with the general decrease of this mark, ~15% of the H3K27me3-marked genes identified in WT seedlings detectably lack the mark in *2h1* plants (Figure S2 and Additional File 1). Spike-in normalized differential analysis of H3K27me3 levels further allowed to estimate that 4317 (~55%) of the H3K27me3-marked genes displayed lower levels of the mark (hypo-marked genes) while only 496 of them showed increased levels (hyper-marked genes) in *2h1* plants (Figure 2A-C, S3 and Additional File 1). Hence, decreased levels of H3K27me3 in *2h1* seedlings formerly identified by immunoblot and immunocytology largely result from a diffuse effect over the majority of the PRC2-targeted genes.

We first envisaged that genes with defective H3K27me3 marking in *2h1* plants might be particularly occupied by H1 in WT plants. This is apparently not the case given that, on the contrary, the set of hypo-marked genes tends to display lower H1 levels in WT plants than unaffected genes (Figure 2D). According to these observations, variations in local H1 abundance may contribute to determine PRC2 marking but the mechanism remained unclear. We therefore explored whether a sequence-dependent mechanism could underlie the differential influence of H1 by undertaking an agnostic search of sequence motifs in the promoters of the gene sets with different H3K27me3 marking. This did not reveal any over-represented sequences in the promoters of hypo-marked genes as compared to all other H3K27me3-marked genes. In contrast, we found three prevalent sequences among the set of 496 hyper-marked genes (Figure S4). Among them, monomeric AAACCCTA telomeric motifs, referred to as *telobox* regulatory elements (Regad et al., 1994; Tremousaygue et al., 1999) can serve as *Polycomb* Response Elements in *Arabidopsis* (Xiao et al., 2017; Zhou et al., 2018). Globally, 87 of the 496 hyper-marked genes contain one or more *telobox* in their promoter (Additional File 1).

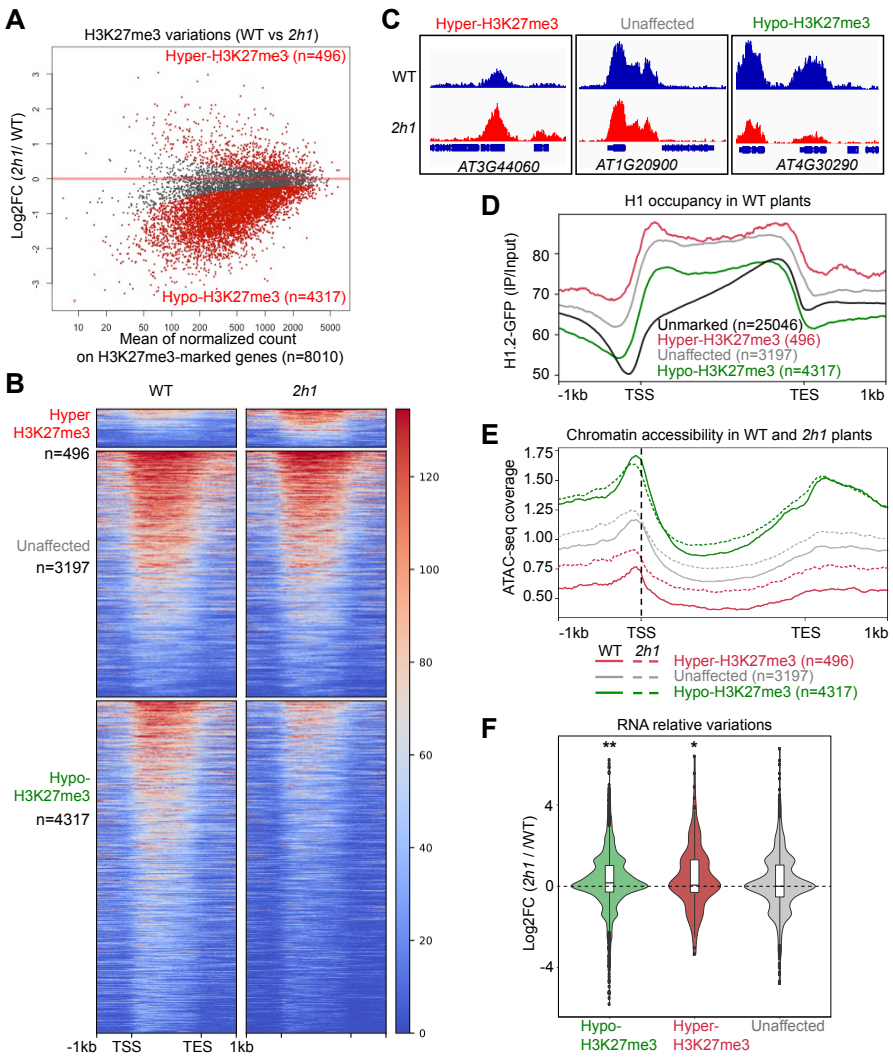


Figure 2. H1 influences H3K27me3 marking, chromatin accessibility and expression of PRC2-targeted genes.

A. Identification of differentially marked genes using spike-in normalized DESeq2 analysis shows that *H1* loss-of-function triggers low H3K27me3 levels over a majority of the PRC2 targeted genes. All genes displaying a H3K27me3-enriched domain in WT or *2h1* plants (according to MACS2 peak detection, see Methods) are individually shown as dots. Red dots represent the differentially marked genes (FDR < 0.01). **B.** H3K27me3 profiles along all genes marked in WT or *2h1* plants. Genes are grouped according to H3K27me3 levels after spike-in normalization and ranked in each group according to mean H3K27me3 levels. **C.** H3K27me3 profile on representative genes of the three categories identified in (A) exemplifying the general tendency of PRC2-targeted genes to keep a weak H3K27me3 domain in *2h1* plants. **D.** H1.2-GFP profile on the same three gene sets. **E.** ATAC-seq analysis of the three H3K27me3-marked gene groups identified in (A). Mean read coverage (reads per million) is used as a proxy of chromatin accessibility. **F.** Transcript level variations in the same three gene sets. The values represent log2 fold changes between *2h1* and WT. The embedded box plots display the median while lower and upper hinges correspond to the first and third quartiles. * indicates a p-value < 5% and ** < 1% according to a Student's t-test.

Having identified that H1 is required for H3K27me3 marking at thousands of PRC2-targeted genes, we examined whether differentially affected genes had distinct chromatin accessibility in WT plants. H3K27me3 hypo-marked genes, with moderate H1 occupancy, were globally more accessible than unaffected or hyper-marked genes, which were particularly H1-rich (Figure 2D) and very poorly accessible (Figure 2E). To test whether their

chromatin accessibility depends on H1, we used *2h1* plants analyzed in parallel with ATAC-seq. This revealed that hypo-marked gene bodies are slightly more accessible in mutant than in WT plants (Figure 2E), thereby correlating with the loss H1 and the decrease of H3K27me3. However, accessibility of unaffected but also of hyper-marked gene sets was also increased in *2h1* plants. Hence, in the latter case, H1 depletion is followed by an increase of chromatin accessibility despite a clear gain in H3K27me3. This discrepancy suggests that H1, more than changes in H3K27me3, influences the accessibility of PRC2-targeted genes.

H1 contributes to define expression of PRC2 targeted genes

To get more insights about the biological role of H1 on PRC2-targeted genes, we analyzed the transcript levels of H3K27me3-marked genes in *2h1* plants and examined the functional categorization of the misregulated ones. Similarly to previous reports (Choi et al., 2019; Rutowicz et al., 2019), our RNA-seq analysis identified that *H1* mutations trigger minor gene expression defects (348 misregulated genes – Additional File 2). Nonetheless, focusing on the differentially marked genes showed that the set of 4317 H3K27me3 hypo-marked genes had globally higher transcript levels in *2h1* than in WT plants (Figure 2F). Gene ontology analysis of this gene repertoire identified an over-representation of genes involved in transcriptional regulation, meristem maintenance, cell wall organization and vascular development (Figure S5A), which are consistent with the repression of these biological processes by PRC2 (de Lucas et al., 2016) and with the subtle PRC2-like phenotypes found in *H1* mutant plants (Rutowicz et al., 2019). Contrasting with the hypo-marked genes, a large proportion of the hyper-marked genes displayed TE-like features, as 25% (124/496) of them overlaps with a TE annotation or is annotated as transposon gene (Figure S5B, Additional File 1).

Collectively, our first analyses identified a positive influence of H1 on H3K27me3 enrichment while restricting chromatin accessibility on a majority of PRC2-targeted genes, which may jointly influence their expression. This trend is contrasted by a reciprocal effect in preventing H3K27me3 accumulation at a minority of H1-rich genes that unexpectedly display frequent TE features.

H1 prevents H3K27me3 invasion over a specific family of heterochromatic repeats

Considering the *de novo* marking of H3K27me3 on several TE-related genes in *2h1* plants, we extended our analysis to all known *Arabidopsis* TEs. This revealed that 1066 TEs are newly marked by H3K27me3 in *2h1* plants, most frequently over their entire length, thereby excluding *a priori* the possibility of a spreading from neighboring genes. We clustered the newly marked TEs into two main groups, *TE cluster 1* (n=216) with H3K27me3 enrichment exceeding almost 2-fold those of PRC2-targeted genes, and *TE cluster 2* (n=850) with milder H3K27me3 enrichment (Figure 3A). While *TE cluster 2* is composed of a representative variety of TE super-families, *TE cluster 1* mostly consists of "*DNA/Others*" repeat annotations prevalently consisting in *ATREP18* elements (Figure 3B). This corresponds to a very high over-representation as *TE cluster 1* and *2* comprise 189 and 47, respectively, of the 391 *ATREP18s Arabidopsis* elements, including many of the longest units (Figure S6).

Comparison of H3K27me3 variations with H1 occupancy showed that *ATREP18* elements also stand out from the general population of TEs by their outstanding gain in H3K27me3 in *2h1* plants (red dots in Figure 3C). *TE cluster 1* and more generally *ATREP18* elements not only display a heterochromatic properties with H3K9me2 marking (Figure S7A-E), elevated nucleosome and H1 occupancies (Figure 3D), but they also have extremely low chromatin accessibility as compared to the majority of TEs (Figure 3E and S7F). Taken together, these observations indicate that H1 has a repressive effect on H3K27me3 over a set of H1-rich, heterochromatic, and highly compacted repeats, which contrasts with its positive influence on H3K27me3 marking over thousands of PRC2-targeted genes.

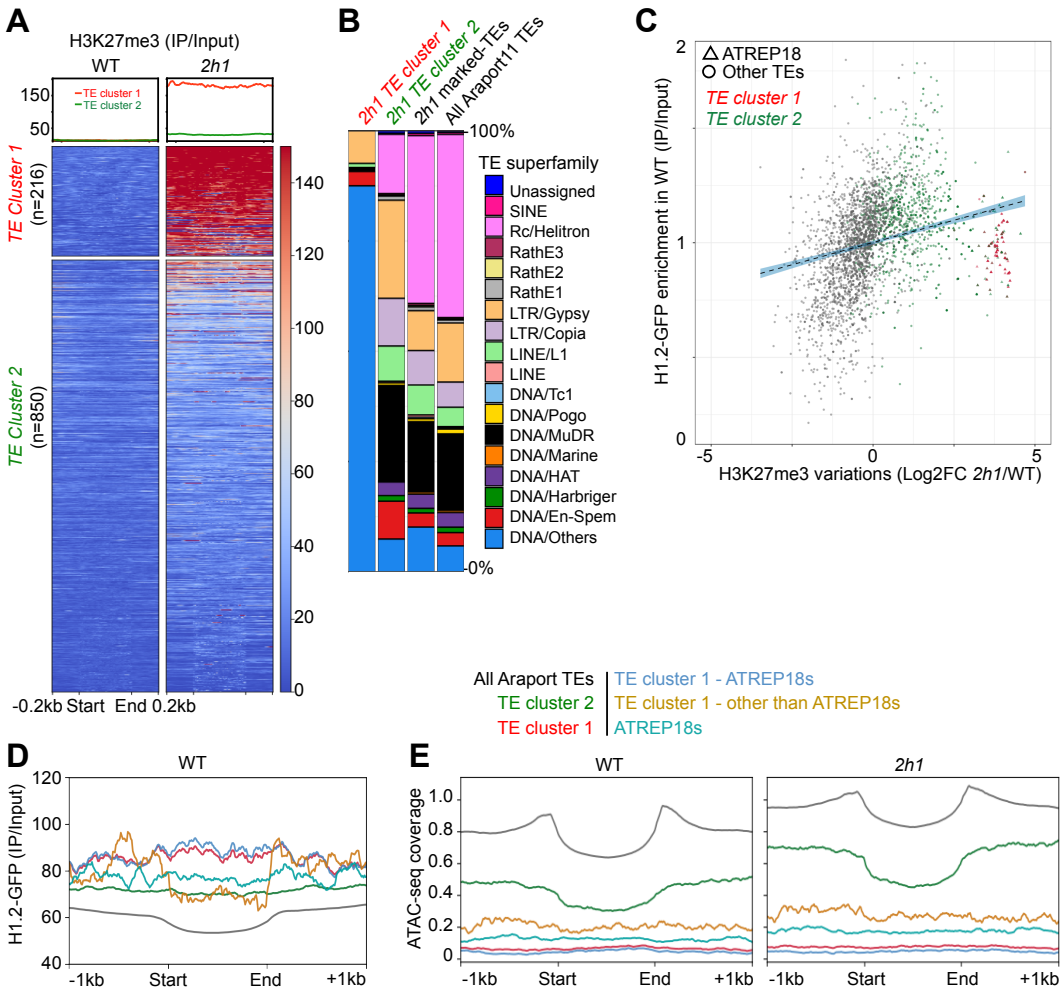


Figure 3. H1 is required to avoid elevated H3K27me3 levels on a specific set of repeated elements. A. H327me3 profiles over all *TE Cluster 1* and *2* elements. Hyper-marked TEs are grouped according to H3K27me3 levels after spike-in normalization and ranked in each group according to mean H3K27me3 levels. **B.** The bar chart represents the relative TE superfamily composition of the indicated sets of TEs. *TE cluster 1* comprises a strong over-representation of "DNA/Others" annotations while *TE cluster 2* elements correspond to a wide variety of TE super-families. **C.** *ATREP18* elements are visible as an outlier group of TEs characterized by strong H3K27me3 enrichment in *2h1* plants. The scatter-plot represents all TE elements marked by H3K27me3 in *2h1* plants. TEs are distributed according to H1.2-GFP level along TE length in the y-axis and H3K27me3 variations (*2h1*/WT) over TE length (Log_2FC) in the x-axis. **D.** *TE Cluster 1* elements display high H1 occupancy. The metagene plot represents mean H1.2-GFP level over the indicated repertoire of TEs and repeats. **E.** Low chromatin accessibility of *ATREP18s* and more generally of *TE cluster 1* elements is conserved in *2h1* nuclei. Mean read coverage in ATAC-seq analysis (reads per million) is used as a measure of chromatin accessibility of the indicated repertoire of TEs and repeats.

While TEs are on average more accessible in *H1* mutant plants (Choi et al., 2019; Rutowicz et al., 2019), ATAC-seq analysis of *TE cluster 1* and *ATREP18* repeats showed that chromatin of these elements remained poorly accessible in *2h1* plants (Figure 3E and S6B). This observation indicates that chromatin inaccessibility of *ATREP18* elements is either H1-independent or that H1 depletion is compensated by other mechanisms such as H3K27me3 enrichment.

Large blocks of heterochromatic interstitial telomeric repeats gain H3K27me3 in *2h1* plants

To explore how H1 selectively contributes to repress H3K27me3 marking on *TE cluster 1*, we first envisaged the H3K27me1 heterochromatic mark as a potential substrate for H3-Lys27 trimethylation. Analysis of public datasets (Ma et al., 2018) showed however that, in contrast to other TEs, H3K27me1 is not abundant on *TE cluster 1* nor on *ATREP18* elements (Figure S7D). We further examined the epigenomic profile of loss-of-function plants for the three major histone H3K27me3 demethylases EARLY FLOWERING 6, RELATIVE OF ELF 6, and JUMANJI 13 (Yan et al., 2018). This revealed no evidence for active H3K27me3 removal on *TE cluster 1* elements in WT plants (Figure S8), altogether ruling out a specific role for H1 in promoting transitions conversion of H3K27me1 into H3K27me3 or in H3K27me3 erasure at these loci.

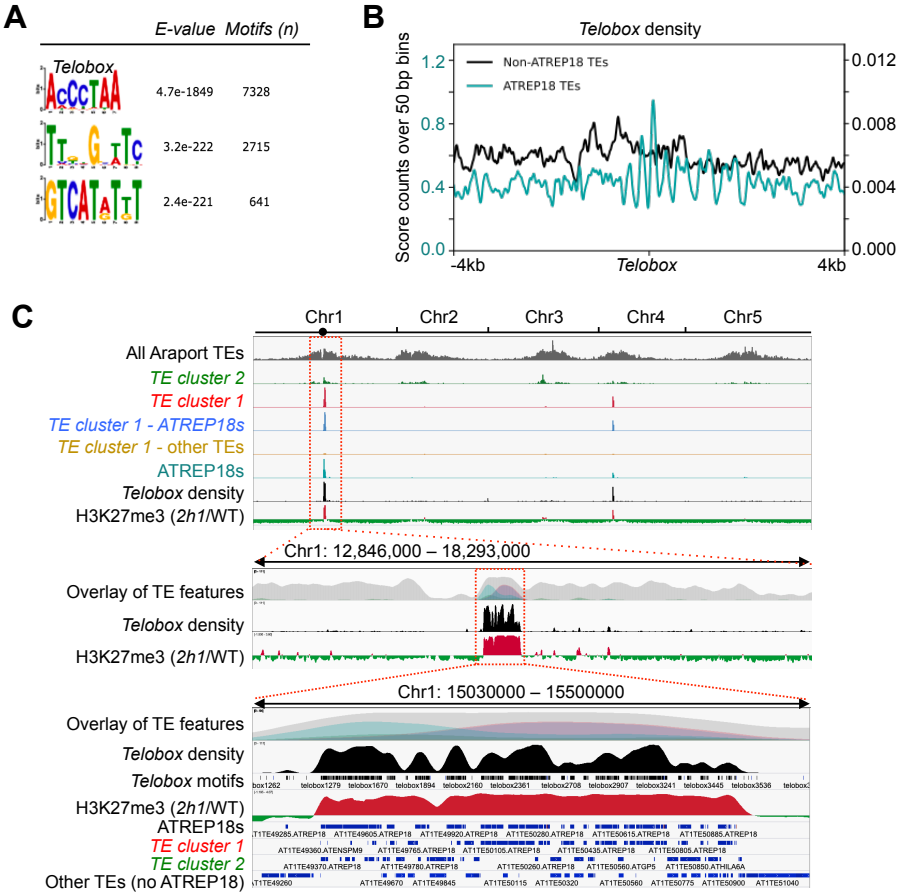


Figure 4. Two pericentromeric ITR blocks, altogether spanning ~430 kb, are subject to massive H3K27me3 enrichment in *2h1* plants. A. Prevalent sequence motif over-represented in *TE cluster 1* elements (E-value<1e-220). E-values were calculated against all Araport11 TE sequences. **B.** *ATREP18* elements display outstanding density and a distinct pattern of *telobox* motifs as compared to other TEs. The plot represents the density of perfect *telobox* sequence motifs over all *ATREP18s* or over all other TEs as 50bp bins surrounding each *telobox* sequence. Note the different colored scales of each TE set. **C.** Genomic distribution and H3K27me3 profiles of *ATREP18*, *TE Cluster 1* and *TE Cluster 2* elements over ITR-1R. The magnification over chromosome 1 allows visualizing a sharp overlap between *2h1*-specific H3K27me3 enrichment and the *telobox*-rich region that we defined as ITR-1R. Complementary magnifications over ITR-4L and over several examples of dispersed elements of *TE cluster 2* are shown in Figure S11.

We subsequently searched for over-represented DNA motifs in *TE cluster 1-2* elements. Out of the three 5-9 bp motifs identified, the most significantly enriched sequence corresponds to the *telobox* (ACCCTAA) motif

(Figure 4A), recalling the same finding for many TE-related genes gaining H3K27me3 in *2h1* plants (Figure S4). In total, 7328 *teloboxes* were found distributed within 195 of the 216 *TE cluster 1* elements (90%). Similar motif search was performed on the 391 *ATREP18s* annotated elements and unveiled more than ten thousand perfect *telobox* motifs, which appeared to be frequently organized as small clusters (Figure S9). Vice versa, profiling of *telobox* motifs in the entire genome revealed their ~100-fold over-representation in *ATREP18* elements as compared to other TEs (Figure 4B). *ATREP18s* display neither typical TE functional features nor predictable protein-coding region (Figure S10A), are mostly oriented on the minus DNA strand (Figure S10B), and tend to be located in close vicinity, nearly 90 % of them being positioned within 1kb of each other (Figure S10C). Consistently with this spatial proximity, analysis of *ATREP18* chromosomal distribution revealed their concentration within two regions of ~355 and ~72 kb on chromosomes 1 and 4, respectively, which both colocalize remarkably well with H3K27me3 enrichment in *2h1* plants (Figure 4C and S11).

These specific features lead us to envisage these large blocks as two of the nine genome loci previously considered as Interstitial Telomeric Repeats or ITRs (Uchida et al., 2002; Vannier et al., 2009), which also comprise a mix of perfect and degenerated telomeric repeats (Schrumpfová et al., 2019). About 95 % of *TE Cluster 1* elements are indeed contained within the two largest ITR blocks identified within pericentromeric regions of chromosomes 1R and 4L (ITR-1R and ITR-4L) (Figure 4C). H3K27me3 ectopic deposition was also found on interspersed *TE cluster 2* elements located in the pericentromeres of the five chromosomes outside from these two ITR blocks (Figure S11B), but our main conclusion is that H1 abundantly occupies two large blocks of pericentromeric ITRs where it prevents H3K27me3 marking.

H1 excludes the GH1-containing TRB1 protein from ITR blocks

With the aim of assessing the molecular mechanisms that selectively drives PRC2 activity on ITRs upon H1 depletion, we envisioned that Telomere Repeat Binding (TRB) proteins might have a prominent role. The TRB1-TRB3 founding members of this plant-specific family of GH1-containing Single-Myb-histone proteins constitute part of the telomere nucleoprotein structure required for telomere length maintenance (Schrumpfová et al., 2014). Their Myb domain has strong affinity to the G-rich strand of *telobox* DNA motifs (Mozgová et al., 2008; Schrumpfová et al., 2004, 2014) and allows for a general role as transcriptional regulators (Schrumpfová et al., 2016; Zhou et al., 2016) recently implicated in PRC2 complex recruitment onto gene promoters (Xiao et al., 2017; Zhou et al., 2018). Interestingly, despite low protein sequence similarity with H1 proteins (14±2%; Figure S12), TRBs display a typical GH1 domain (Charbonnel et al., 2018; Kotliński et al., 2017). Hence, we reasoned that reciprocal binding of H1 and TRB proteins to chromatin might modulate PRC2 recruitment on ITRs.

To explore this possibility, we first examined the genomic distribution of TRB1 with regard to H1 and *telobox* motifs using available ChIP-seq data (Schrumpfová et al., 2016). After confirming that TRB1 peaks found over genes are centered on *telobox* motifs, and that H1 is globally enriched over *TE Cluster 1* elements, this showed that H1 level appears to be anti-correlated with *telobox* positioning and with TRB1 in these elements (Figure 5A and S13). To assess whether this apparent antagonism is a general property, we examined H1 occupancy over all TRB1 genome binding sites (Schrumpfová et al., 2016) and found an inverse correlation between H1 and TRB1 both on genes and TEs (Figure 5B). Consistent with the hypothesis that H1 prevents PRC2 recruitment over telomeric repeats, these observations hint at a reciprocal relation for DNA binding of H1 and TRB1 over *telobox*-rich regions, including on *TE Cluster 1*.

To better resolve this apparent mutual exclusion and link it to nucleosome core and linker DNA positioning, we took advantage of the well-positioned nucleosome (WPN) coordinates previously defined using MNase-seq (Lyons and Zilberman, 2017) and plotted the profiles of H1, TRB1, *telobox* motifs and nucleosome occupancy over all WPNs commonly present in WT and *2h1* plant lines. While H1 was expectedly enriched over DNA linker regions, TRB1 tends to exhibit a much broader distribution spanning 4-5 nucleosomes. TRB1 profile

nonetheless displays a periodic pattern matching linker DNA regions, and, moreover, *telobox* distribution itself is sharply coincident with regions serving as linker DNA (Figure 5C). Given this unexpected preferential distribution, *telobox* motifs apparently play a functional role in chromatin organization and are subjected to reciprocal binding of H1 and TRB1 as schematized in Figure 5D.

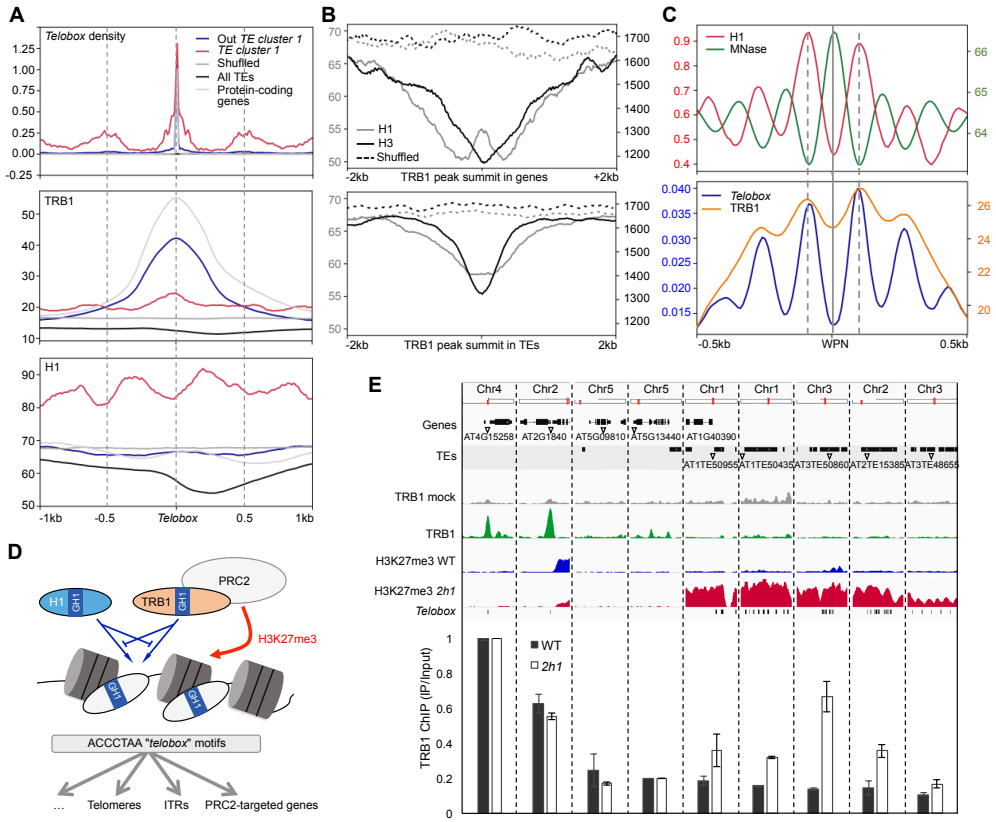


Figure 5. H1 restricts the accessibility of TRB1 to *telobox* loci. A. Global occupancy of TRB1 and H1 at all *telobox* positions in WT plants. TRB1 median peak summit expectedly correlates with the position of *teloboxes* located in protein-coding genes. In contrast, TRB1 is weakly associated to the multiple *telobox* motifs of *TE cluster 1* elements, which are globally H1-rich. While appearing both influenced by *telobox* positioning, H1 and TRB1 patterns appear to be anti-correlated over the *TE cluster 1* elements. **B.** Comparison of H1 and H3 occupancy over all TRB1 peaks unveils a clear depletion of these two histones on TRB1 median peak summits. **C.** H1, TRB1 and *telobox* motifs tend to be distributed over DNA linker regions. The genome-wide profiles of H1, TRB1 and *telobox* sequence motifs were plotted over the coordinates of all *Arabidopsis* well-positioned nucleosomes defined by MNase-seq in Lyons and Zilberman (2017). Analysis of the MNase profile shows the expected linker DNA profile of H1.2-GFP and reveals that median TRB1 peak summits and the density of *telobox* motifs are pronounced on inter-nucleosomal regions. **D.** Working model of possible competitive binding of TRB1 with H1 on linker DNA-localized *telobox* motifs. Following this model, high H1 occupancy on ITRs restricts TRB1 DNA binding and possibly also to other *telobox* loci. In *2h1* plants, increased DNA accessibility would enable the sequence-specific binding of TRB1 to telomeric repeat sequences and recruit PRC2 activity, thereby triggering *de novo* H3K27me3 deposition. **E.** Ectopic enrichment of TRB1 over ITR-1R and several *TE cluster 2* elements in *2h1* plants. Anti-TRB1 ChIP was performed using WT and *2h1* plants before analysis by qPCR using known TRB1 associated genes (*AT4G15258* and *AT2G1840*) as positive controls, and two genes with no TRB1 signal (*AT5G09810* and *AT5G13440*) as negative control. Five loci displaying a clear H3K27me3 enrichment over *telobox* motifs in *TE cluster1* or *2* were tested. Error bars represent standard deviation from two independent biological replicates.

Based on these findings, we envisaged that ectopic H3K27me3 enrichment over *TE cluster 1* and *2* may result from TRB recruitment to *telobox*-rich loci when H1 is depleted. We tested this possibility by performing ChIP analyses of endogenous TRB1 in *2h1* plants. Analysis of two known TRB1 target gene promoters (Schrumpfová et al., 2014; Schrumpfová et al., 2016) and two non-target genes without *telobox* showed the specificity of our assay (Figure 5E top panel). We then tested *telobox*-rich loci subject to ectopic H3K27me3 deposition, two loci in ITR-1R and three interspersed *ATREP18* or *LTR-Gypsy* elements selected from *TE cluster 2* and located on the other chromosomes. Corroborating the ChIP-seq analyses, no significant TRB1 association was detected over these five loci in WT plants but TRB1 enrichment was reproducibly detected in *2h1* samples (Figure 5E bottom panel). These findings support a role for H1 in preventing TRB recruitment on *telobox* containing *TE Cluster 1-2* elements, and provide a plausible mechanism for *de novo* H3K27me3 deposition on ITRs in the absence of H1.

A role for H1 in telomere chromatin composition and organization

Considering their obvious density in telomeric repeats and their propensity to attract TRB proteins (Schrumpfová et al., 2014), we assessed whether, similarly to ITRs, telomeres are subject to H3K27me3 enrichment in *2h1* plants. Because the perfect continuum of terminal telomeric motifs is not suited for quantitative NGS analyses, we examined telomeric H3K27me3 using dot-blot hybridization of ChIP DNA to radioactively labeled concatenated telomeric probes (Adamusová et al., 2020). WT and *2h1* chromatin extracts were immunoprecipitated using either anti-H3K27me3 or anti-H3 antibodies. This led to the estimation that telomeres display an average ~4-fold H3K27me3 enrichment in *2h1* as compared to WT plants, independently from any detectable change in nucleosome occupancy (Figure 6A and S14).

To assess whether H3K27me3 enrichment concerns specific telomeres or affects them generally, we explored the nuclear distribution of this histone mark by immunolabeling combined with telomere Fluorescence In Situ Hybridization (telomere FISH). This complementary approach showed that H3K27me3 signals colocalize with most visible telomeric foci, which are usually distributed around the nucleolus in WT nuclei. Consistent with our dot-blot analysis, H3K27me3 signal intensity at telomere foci was enhanced in the *2h1* line (Figure 6B-C). Moreover, two-to-four telomere foci frequently presented outstanding H3K27me3 signals (Figure 6B). We did not ascertain whether these two atypical FISH signals correspond to telomeres or to cross-hybridizing pericentromeric ITRs from chromosomes 1R and 4L. Their frequent positioning near to the nuclear periphery coincides with the typical localization of pericentromeres and might hint at the latter hypothesis, albeit many telomeric foci are abnormally distant from the nucleolus in *2h1* plants (Figure 6C).

In this analysis we also unexpectedly detected a decreased number of telomeric foci in *2h1* as compared to WT plants (Figure 6C). This cytogenetic pattern may result from defective individualization of the telomeres and to indirect topological alterations leading to their mislocalization in the nuclear space. Collectively, we concluded that H1 does not only prevent excessive accumulation of H3K27me3 over ITRs and telomeres (Figure 6D) but also influences the interphase sub-nuclear organization of chromosomes.

H1 depletion triggers the insulation of the ITRs newly marked by H3K27me3

To gain a more detailed view of chromosome organization defects induced by *H1* loss-of-function and investigate how ITRs could be impacted, we employed *in situ* chromatin conformation capture (Hi-C) on WT and *2h1* nuclei isolated from dissected cotyledons. The tissue homogeneity and high read coverage allowed us to reach a 1 kb resolution (Figure S15). Consistent with previous reports (Feng et al., 2014; Grob et al., 2014; Liu et al., 2016; Moissiard et al., 2012; Sun et al., 2020), WT plants displayed frequent intra-chromosomal interactions within the pericentromeric regions and much less within the chromosomal arms (Figure S16).

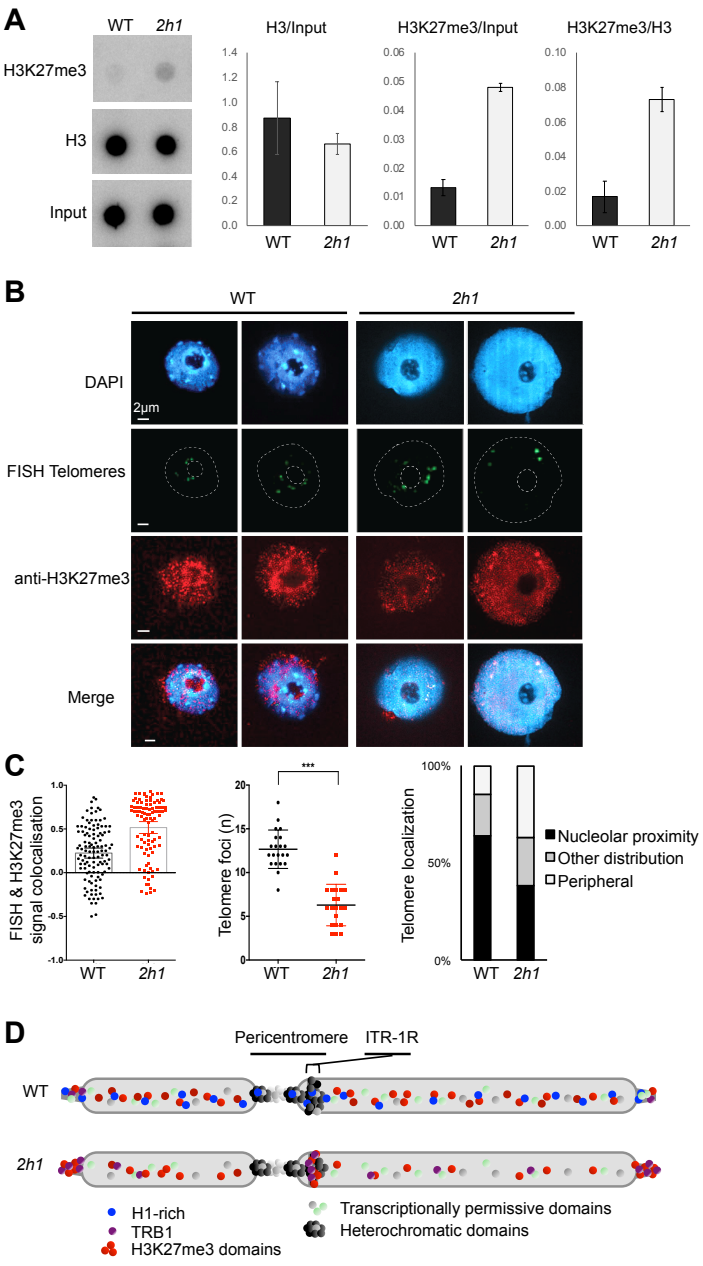


Figure 6. H1 influences H3K27me3 enrichment and the sub-nuclear organization of telomeres. **A.** ChIP analysis of H3K27me3 and histone H3 shows that H3K27me3 signal relative to nucleosome occupancy is enhanced in *2h1* plants. Anti-H3K27me3 ChIPs were followed by dot-blot hybridization with a radioactively labeled telomeric probe in two biologically and technically replicated experiments. The second biological replicate is given in Figure S14. **B.** H3K27me3 immunodetection and telomeric DNA FISH analysis of WT and *2h1* cotyledon nuclei show that multiple telomeric loci are H3K27me3-enriched in *2h1* plants. Nuclei were counterstained with DAPI (blue), telomere FISH signals are colored in green and H3K27me3 immunolabeling signal is colored in red. Images are collapsed Z-stack projections. Bars = 2 µm. **C.** Quantification of sub-nuclear telomeric signal properties. The estimated number and position of telomeric foci was determined in two independent biological replicates (n>20). **D.** Working model. Elevated H1 occupancy on ITRs might represent a safeguarding mechanism to avoid the formation of gigantic H3K27me3-rich domains in chromosome 1 pericentromere and telomeres. In this model, H1 depletion permits TRB-mediated recruitment of PRC2 onto densely repeated telomeric motifs located in the telomeres and on the 355 kb-long ITR-1R. The hypotheses of TRB-mediated PRC2 recruitment over telomeric repeats, as well as a putative titration of PRC2 activity away from protein-coding genes both remain to be tested.

Examination of interaction decay exponents (IDEs) of intra-chromosomal interactions unveiled a small tendency for more short-distance and fewer long-distance interactions in the mutant line (Figure 7A). In line with an impaired chromocenter formation in *2h1* cotyledon nuclei (Choi et al., 2019; Rutowicz et al., 2019), we also detected a strong decrease of intra-pericentromeric chromatin interactions (Figure 7A-B). We then examined ITR-1R and 4L and observed that intra-ITR interactions were strongly enhanced in *2h1* plants, thereby resembling typical topologically associating domains (TADs) (Figure 7C). Hence, defects in chromatin topology of ITR-1R and ITR-4L display an opposite trend to their neighboring pericentromeric environment, which is characterized by extensive relaxation of heterochromatin when H1 is depleted. As highlighted in the bottom panel of Figure 7C, magnification in this transition reveals a sharp correspondence between H3K27me3-enriched ITR-1R borders and the more relaxed surrounding pericentromeric regions in *2h1* plants.

Finally, to assess more broadly the relationships between topological and H3K27me3 defects in *2h1* plants, we plotted all differences in H3K27me3 profiles over the chromosomes and compared them to a similar representation of the Hi-C data (Figure 7D). The comparison sheds light on ITR-1R and ITR-4L as two H3K27me3-enriched genome loci with prominent differences in chromatin interactions in *2h1* chromatin.

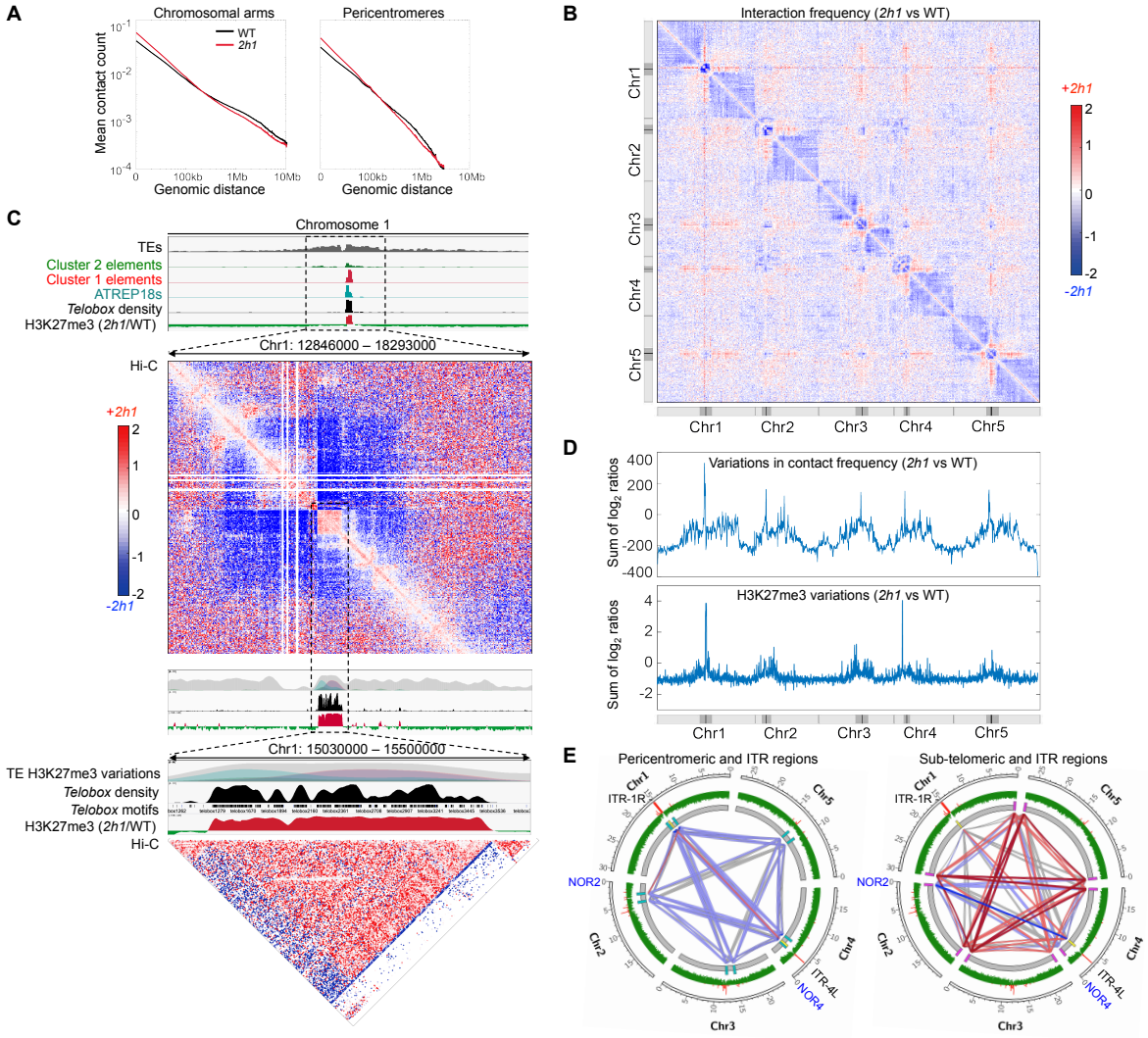


Figure 7. H3K27me3 enrichment on ITR-1R and on the telomeres associates with higher-order topological defects. A. Mean contact count as a function of genomic distance for all chromosomes at a 10 kb resolution (IDE). **B.** Differential interaction frequencies between WT and *2h1* plants. Regions in red have more frequent contacts in *2h1* than in WT plants; regions in blue have an opposite behavior. The Log_2 values of interaction frequency of the five chromosomes in *2h1* versus WT are shown at a 100 kb resolution. The top and side tracks show the positions of centromeres (black) and pericentromeric domains (grey). **C.** In *2h1* mutant plants, a general reduction of intra-pericentromeric interaction frequencies is contrasted by a TAD formation matching the H3K27me3-enriched ITR-1R 355 kb locus. Top panel, location of ITR-1R in chromosome 1. Middle panel, contrasted behaviors of pericentromeric regions surrounding the TAD-like structure formation over ITR-1R in *2h1* plants at a 10 kb resolution. Bottom panel, magnification of the ITR-1R locus at a 2kb resolution shows a sharp correspondence at ITR-1R boundaries between H3K27me3 enrichment and topological changes induced by *H1* loss-of-function. **D.** Comparison of Hi-C and H3K27me3 genome-wide variations highlights the wide impact of ITR-1R and ITR-4L regions on topological changes induced by *H1* loss-of-function. Top panel: sum of all log_2 ratios from Hi-C contact map variations using 100 kb bins. Bottom panel: similar

analysis for H3K27me3 variations using 10 kb bins. **E.** Variations in interaction intensities among the telomere proximal, pericentromeric and ITR regions show that topological interaction patterns are specifically altered by *H1* loss-of-function. Except for the NOR-associated Chr2L and Chr4L regions, interactions of telomeres between them and with the H3K27me3-enriched ITRs tend to be more frequent in the mutant line. Similarly, interaction frequency between the ITRs is increased in the absence of H1. Interestingly, H1 depletion also triggers an outstanding increase of interaction frequency between the NOR2-associated region and ITR-4L. Yellow boxes, ITR regions. Magenta boxes, telomere proximal regions. External green/red track, H3K27me3 variations in *2h1* versus WT plants (Log2 ratio). Dark and pale red linkers, strong or moderate increase in *2h1* plants, respectively. Blue linkers, strong decrease in *2h1* plants. More details on the regions selected and on control loci are given in Figure S17.

H1 impacts long-distance interactions between chromosome ends

Considering that H3K27me3-enriched ITRs and telomeres display altered sub-nuclear positioning in *H1* mutant plants, we examined long-distance interactions among these loci in the Hi-C matrixes. Because telomeres are not included in the *Arabidopsis* reference genome, we used the most terminal 100 kb regions of each chromosome sequence as a proxy to probe the 3D organization of subtelomeric regions. In this analysis, we also considered an internal 100 kb region for each pericentromeric region, ITR-1R and ITR-4L coordinates, as well as several 100 kb regions randomly chosen in distal chromosomal arms. Possibly reflecting the capacity of different centromeres to aggregate within chromocenters, we observed that pericentromeric regions tend to frequently interact with each other as compared to chromosomal arm domains in WT plants (Figure S17A). In agreement with the preferential localization of telomeres at the nuclear periphery and of the centromeres at the nuclear interior (reviewed in Pontvianne and Grob, 2020; Santos et al., 2020), a similar tendency was observed for most of the telomere-proximal regions but not for ITR/telomere interactions (Figure S17A). Two outstanding exceptions were the northern tips of chromosomes 2 and 4 that neighbor the ~4 Mb Nucleolar Organizing Regions (*NOR2* and *NOR4*) (Copenhaver and Pikaard, 1996), which tend to be more isolated than all other telomere proximal regions (Figure S17A).

We then examined the *2h1* line and observed that, excluding the regions linked to *NOR2* and *NOR4*, the frequency of interactions between all the sub-telomeric regions were detectably increased in the absence of H1 (Figure 7E and S17A-C). This observation supports an organizational model in which telomere territories tend to coalesce more frequently in the absence of H1, as hinted by the lower number of telomere FISH signals and the appearance of strong H3K27me3-marked telomeric foci in *2h1* nuclei (Figure 6B). Examination of ITR-1R and ITR-4L also showed a slight tendency of increased association between these two domains, but not with the subtelomeric regions (Figure 7E and S17). Hence, contrasting with its positive impact on intra/inter-pericentromere associations, H1 appears to be required for dampening long-distance associations between ITRs or telomeres in addition to preventing H3K27me3 enrichment on these loci.

Discussion

We have reported that H1 is highly abundant on the gene repertoire targeted by PRC2 where it contributes to efficient H3K27me3 marking and to diminish chromatin accessibility. Vice versa, our observations also hint at a repressive role for H1 in H3K27me3 enrichment on ITRs and telomeres. The large scale on which these two antagonistic patterns are observed sheds light on the existence of strong mechanistic links between H1 and *Polycomb*-based regulation in *Arabidopsis*, two main actors in the instruction of DNA accessibility.

Functional interplays between H1 and H3K27me3 for the regulation of protein-coding genes

We have not been able to identify any sequence specificity for H1-mediated H3K27me3 gene marking. Although 4317 of the PRC2-targeted genes were hypo-marked in *2h1* plants, the large majority of the H3K27me3 peaks

identified in WT plants were still present. This suggests that H1 has a general influence on H3K27me3 deposition/maintenance or spreading but is not necessarily mandatory for nucleation of H3K27me3. While in most instances H1 depletion resulted in an increase of chromatin accessibility of PRC2-targeted genes, its impact on gene expression was apparently more related to variations in H3K27me3 marking. Hence, consistently with the subtle phenotypes of *H1* mutant plants (Rutowicz et al., 2019), part of the defects in gene activity observed after H1 depletion might result from indirect consequences through PRC2 function.

Contrasting with the positive influence of H1 on H3K27me3 marking over thousands of genes, we also identified an opposite role of H1 in limiting H3K27me3 deposition over a minority (496) of protein-coding genes, which often contained a *telobox* sequence motif in their promoter and frequently display TE features. In *2h1* mutant plants these genes are heavily marked by H3K27me3 and still efficiently repressed while their chromatin generally remained poorly accessible. These reciprocal effects underlie a dual influence of H1 on the expression and chromatin status of distinct loci.

Sequence-specific repression of H3K27me3 enrichment over telomeric repeats

Other than protein-coding genes, we noted that H1 massively minimizes H3K27me3 abundance over two large ITRs and on the telomeres. These ITRs are mainly composed of *ATREP18* heterochromatic elements that lack an open reading frame, contain repeated clusters of *telobox* motifs, and are essentially crowded together in the pericentromeres of chromosomes 1 and 4. The identification of such an antagonistic effect of H1 on H3K27me3 over thousands of protein-coding genes on the one hand and over telomeric repeats on the other hand provides an explanation for our former observation that most of H3K27me3 sub-nuclear signals are low in *H1* loss-of-function plants whilst, intriguingly, a few foci of undetermined nature remain enriched (Rutowicz et al., 2019). Here, cytogenetic, epigenomic and Hi-C analyses revealed that these enriched foci correspond to telomeric repeats, most probably to ITR-1R and ITR-4L.

Our analyses of nuclear organization refine the recent observation that chromocenter formation is impaired in *2h1* leaf and cotyledon nuclei (Choi et al., 2019; He et al., 2019; Rutowicz et al., 2019), a defect that commonly reflects the spatial dispersion of pericentromeres within the nuclear space (Fransz et al., 2002). Hi-C analyses of *2h1* nuclei first identified a reduced frequency of chromatin interactions within and among the pericentromeres, which is a typical feature of *Arabidopsis* mutant plants affected in chromocenter formation (Feng et al., 2014; Grob et al., 2014; Moissiard et al., 2012) or when chromocenters are decompacted in response to environmental stress (Sun et al., 2020). Hi-C analyses further revealed a more complex picture in which ITRs embedded within the pericentromeres of chromosomes 1 and 4 escape the surrounding relaxation of heterochromatin induced by H1 depletion. This local contrast sheds light on unexpected differences of H1-mediated regulation of chromosome topology between adjacent heterochromatic domains. As identified using ATAC-seq and Hi-C analyses, H3K27me3 enrichment as a consequence of H1 depletion might underlie the maintenance of extremely lowly accessible ITR chromatin while other heterochromatic regions tend to become more accessible as a consequence of H1 depletion.

H1 may minimize PRC2 activity on telomeric repeats through competitive binding with TRB proteins

Given their capacity to recruit the two somatic PRC2 methyltransferases CURLY-LEAF (CLF) and SWINGER (SWN) (Xiao et al., 2017; Zhou et al., 2018), TRBs represent excellent candidates for a sequence-specific regulation of H3K27me3 on interstitial telomeric repeats and possibly also on telomeres. We observed that TRB1 ectopically associates with ITRs and other *telobox*-rich elements in *2h1* plants, indicating that elevated H1 incorporation on these loci indeed limits TRB1 binding. Based on these findings, we propose that H1 binding to linker DNA restricts TRB1 accessibility to telomeric motifs acting as *PREs* (Figure 5D). Following this model, the strong affinity of TRB1

Myb domain for *telobox* DNA motifs (Mozgová et al., 2008; Schrumpfová et al., 2016) and its highly dynamic association with chromatin (Dvořáčková et al., 2010) may explain TRB1 recruitment to ITRs when H1 is depleted. Ectopic enrichment of TRB1 and H3K27me3 on ITR regions in *2h1* plants may consequently result from increased accessibility of the repeated *telobox* sequences otherwise masked by H1 in WT plants.

Waiting for a mechanistic assessment of the relative affinity of H1 and TRB1 to *telobox* elements in a chromatin context, several other observations support these hypotheses. Firstly, on a genome-wide scale, H1 and TRB1 occupancies are most frequently negatively correlated. Secondly, analysis of nucleosome positioning showed that *telobox* motifs are preferentially situated in linker DNA where TRB1 pattern is also pronounced; hence competition with H1 might occur on linker DNA. The remarkable pattern of *telobox* positioning in linker DNA further suggests a capacity of this short sequence to influence chromatin organization, possibly by repelling nucleosomes away from *telobox* motifs.

Future studies may determine whether H3K27me3 enrichment on telomeric repeats directly relies on PRC2 recruitment by TRB proteins, as recently shown for a gene reporter system (Zhou et al., 2018), or rather implicate other chromatin modifiers influencing H3K27me3. *Arabidopsis* LIKE-HETEROCHROMATIN 1 (LHP1), a PRC1 subunit acting as chromatin reader of H3K27me3 in plants (Turck et al., 2007), indeed prevents TRB1 enrichment on PRC2 target genes displaying individual *telobox* motifs (Zhou el al., 2016). Our model further echoes the recent observation that GH1-containing High Mobility Group A1 (GH1-HMGA1), is present at *Arabidopsis* telomeres and also minimizes H1 incorporation in chromatin (Charbonnel et al., 2018; Kotliński et al., 2017). Hence, H1 chromatin incorporation might act in competition with several GH1 domain-containing proteins, such as TRB1 and GH1-HMGA1. Interestingly, HMG but also other proteins bind to DNA antagonistically to H1 in mammals (Catez et al., 2004; Krishnakumar et al., 2008). In *Arabidopsis*, the repertoire of 15 GH1 domain-containing proteins (Charbonnel et al., 2018; Kotliński et al., 2017) may allow multiple combinations of antagonistic or cooperative associations over the genome.

Functional implication of H1 on telomeric chromatin structure

The nature of plant ITR and telomere chromatin has long remained enigmatic (Achrem et al., 2020; Dvořáčková et al., 2015). Owing to their repetitive nature and to their sequence similarity with ITRs, telomeres are recalcitrant to sequencing-based technologies (Fojtová and Fajkus, 2014; Majerová et al., 2014; Vaquero-Sedas et al., 2011, 2012; Vega-Vaquero et al., 2016). ChIP dot-blot analyses indicate a dominance of H3K9me2 over H3K27me3 histone marks, but some of the telomere regions also display H3K4me2 and H3K4me3 euchromatic marks (Adamusová et al., 2020; Grafi et al., 2007; Vaquero-Sedas et al., 2011). Here, combining ChIP-seq with telomeric probes used in ChIP dot-blot and *in situ* immunolocalization leads us to conclude that H1 does not only prevent H3K27me3 marking of pericentromeric ITRs but is also required to moderate by 2-to-4 fold its accumulation on several individual telomeres. Yet, our analyses did not allow assessing the precise distribution of HK27me3 enrichment along the telomeres, especially if considering their mosaic chromatin status.

Indeed, *Arabidopsis* telomeres are thought to be constituted by chromatin segments with distinct nucleosome repeat length (NRL), one of them being characterized by an average NRL of 150bp (Ascenzi and Gantt, 1999), which is much shorter than the 189 bp size estimated for H1-rich TEs (Choi et al., 2019). Such a small DNA linker size (e.g., 5 bp) is seemingly incompatible with H1 ability to incorporate into chromatin, a process that protects about 20 bp of DNA *in vitro* (Simpson, 1978). Consistently, H1 has been proposed to be under-represented at telomeres in plants (Ascenzi and Gantt, 1999; Fajkus et al., 1995) as in mammals (Achrem et al., 2020; Déjardin and Kingston, 2009; Galati et al., 2013; Makarov et al., 1993). This could explain the short NRL of *Arabidopsis* and human telomeres (Ascenzi and Gantt, 1999; Lejnine et al., 1995) and led to the interpretation that H1-free telomere chromatin segments could display a columnar structure in which nucleosome arrays are stabilized be stacking interactions mediated by the histone octamers themselves (Fajkus and Trifonov, 2001).

Consequently, the existence of distinct chromatin states on *Arabidopsis* telomeres needs to be explored in more details to establish whether the repressive influence of H1 on PRC2 activity is a general property of telomeres or rather impacts only specific segments.

H1 promotes telomere individualization in the nuclear space

We observed that H1 depletion provokes a reduction in the proportion of telomeric foci located nearby the nucleolus and in their total number in the nucleus. Using Hi-C, we could attribute this apparent defect in telomere individualization to more frequent inter-chromosomal interactions at many telomere proximal regions used as a proxy for telomere Hi-C analysis. Preferential positioning of telomeres around the nucleolus and of the centromeres near to the nuclear periphery being an important organizational principle of *Arabidopsis* chromosome territories (reviewed in Pontvianne and Grob, 2020; Santos et al., 2020), H1 appears to be a crucial determinant of *Arabidopsis* interphase nuclear organization.

While both PRC2 and PRC1 participate in defining *Arabidopsis* genome topology (Feng et al., 2014; Veluchamy et al., 2016), H3K27me3 is favored among the long-distance interacting gene promoters (Liu et al., 2016). This has led to the proposal that, as in animals, this mark could contribute to shape *Arabidopsis* chromosomal organization, possibly through the formation of *Polycomb* subnuclear bodies (Liu et al., 2016). Here we mostly focused on large structural components of the genome such as the telomeres, the pericentromeres and ITR regions but H1 depletion also triggers flagrant higher-order topological changes of highly active (transcribed) genome compartments in human embryonic stem pluripotent cells (Geeven et al., 2015). Future studies might help to determine if the broad changes in the H3K27me3 landscape contribute to the chromosome organization disorders observed in *2h1* nuclei, possibly acting in combination with the intrinsic consequences of H1 depletion on chromatin compaction.

A putative role for H1 in modulating H3K27me3 homeostasis between protein-coding genes and large blocks of telomeric repeats

Arabidopsis telomeres span 2 to 5 kb at the end of each chromosome (Fitzgerald et al., 1999; Richards and Ausubel, 1988), a much shorter size than ITR-1R and ITR-4L that bear ~3000 perfect telomeric motifs representing ~107kb and many degenerated ones over ~430kb. The cumulated presence of *teloboxes* in ITR-1R and ITR-4L therefore represents at least 2-fold longer regions than all telomeres considered together, altogether forming an immense potential reservoir of PRC2 targets (Figure 6D). Based on our findings, H1 apparently represents a safeguarding mechanism to avoid the formation of gigantic H3K27me3-rich blocks in the pericentromeric ITRs of chromosomes 1 and 4, on a scale that may eventually tether many PRC2 complexes away from protein-coding genes. In *Neurospora crassa*, artificial introduction of an array of $(TTAGGG)_{17}$ telomere repeats in interstitial sites of the genome triggers the formation of a large block (~225 kb) of H3K27me2/3-rich chromatin (Jamieson et al., 2018). This example and our findings illustrate the intrinsic appetence of telomeric motifs for H3K27 trimethylation in multiple systems.

Methods

Plant lines and growth conditions

The *h1.1 h1.2* (*2h1*) *Arabidopsis* mutant line and the transgenic *pH1.2::H1.2-GFP* line (all in the Col-0 background) were kindly provided by Dr. Kinga Rutowicz (Rutowicz et al., 2015). Seeds were surface-sterilized, plated on half strength Murashige and Skoog (MS) medium with 0.9% agar and 0.5% sugar and cultivated under long-day

(16h/8h) at 23/19°C light/dark photoperiod (100 µmol.m^{-2}.s^{-1}) for 5 days unless otherwise stated. Cotyledons, when used, were manually dissected under a stereomicroscope.

Immuno-FISH

After fixation in PFA 4% in 1XPME, cotyledons of 7-day-old seedlings were chopped directly in 1% cellulase, 1% pectolyase, and 0.5% cytohelicase in 1X PME and incubated 15 min at room temperature. Nucleus suspensions were transferred to poly-Lys-coated slides. One volume of 1% lipsol in 1X PME was added to the mixture and spread on the slide. Then, 1 volume of 4% paraformaldehyde in 1X PME was added and slides were dried at room temperature. Immunodetection and FISH were conducted as described previously (Charbonnel et al., 2018) using the following antibodies: rabbit H3K27me3 (#07-449 - Merck) diluted 1:200, Goat biotin anti Rabbit IgG (#65-6140 - ThermoFisher) 1:500, mouse anti-digoxigenin (#11333062910 -ROCHE) 1:125, rat anti-mouse FITC (#rmg101 - Invitrogen) at 1:500, goat Alexa 488 anti-rabbit IgG (#A11008 – Invitrogen) at 1:100, mouse Cy3 anti-biotin antibody (#C5585 - Sigma) at 1:1000. Acquisitions were performed on a structured illumination (pseudo-confocal) imaging system (ApoTome AxioImager M2; Zeiss) and processed using a deconvolution module (regularized inverse filter algorithm). The colocalization was analyzed via the colocalization module of the ZEN software using the uncollapsed Z-stack files. To test for signal colocalization, the range of Pearson correlation coefficient of H3K27m3 vs telomeric FISH signals were calculated with the colocalization module of the ZEN software using Z-stack files. Foci with coefficients superior to 0.5 were considered as being colocalized.

ATAC-seq

Nuclei were isolated from 200 cotyledons and purified using a two-layer Percoll gradient at 3000 g before staining with 0.5 µM DAPI and sorting by FANS according to their ploidy levels using a MoFlo Astrios EQ Cell Sorter (Beckman Culture). For each sample, 20000 sorted 4C nuclei were collected separately in PBS buffer and centrifuged at 3000 g and 4 °C for 5 min. The nuclei were resuspended in 20 µl Tn5 transposase reaction buffer (Illumina). After tagmentation, DNA was purified using the MinElute PCR Purification Kit (Qiagen) and amplified with Nextera index oligos (Illumina). A size selection was performed with AMPure® XP beads (Beckman Coulter) to collect library molecules longer than 150 bp. DNA libraries were sequenced by Beijing Genomics Institute (BGI Group, Hong-Kong) using the DNA Nanoballs (DNB™) DNBseq in a 65 bp paired-end mode. Raw ATAC-seq data were treated using the custom-designed ASAP (ATAC-Seq data Analysis Pipeline; https://github.com/akramdi/ASAP) pipeline. Mapping was performed using Bowtie2 v.2.3.2 (Langmead and Salzberg, 2012) with parameters --very-sensitive -X 2000. Mapped reads with MAPQ<10, duplicate pairs, and reads mapping to the mitochondrial genome as well as repetitive regions giving aberrant signals (Quadrana et al., 2016) were filtered out. Concordant read pairs were selected and shifted as previously described by 4 bp (Schep et al., 2015). Peak calling was performed using MACS2 (Zhang et al., 2008) using broad mode and the following parameters: --nomodel --shift -50 --extsize 100. Heatmaps and metaplots were produced from depth-normalized read coverage (read per million) using the Deeptools suite (Ramírez et al., 2016).

***In situ* Hi-C**

Hi-C was performed as in Grob et al. (2014) using seedlings crosslinked in 10 mM Potassium Phosphate pH 7.0, 50 mM Sodium Chloride, 0.1 M sucrose with 4 % (v/v) formaldehyde. Crosslinking was stopped by transferring seedling to a new tube containing 30mL of 0.15 M glycine. After rinsing and dissection, 1000 cotyledons were flash-frozen in liquid nitrogen and grinded using a Tissue Lyser (Qiagen). All sample were adjusted to 4 ml using NIB buffer (20 mM Hepes pH7.8, 0.25 M Sucrose, 1 mM MgCl2, 0.5 mM KCl, 40 % v/v Glycerol, 1 % Triton X-100) and homogenized on ice using a Douncer homogenizer. Nuclei were pelleted by centrifugation and resuspended in the

DpnII digestion buffer (10 mM MgCl2, 1 mM DTT, 100 mM NaCl, 50 mM Bis-Tris-HCl, pH 6.0) before adding SDS to a final concentration of 0.5 % (v/v). SDS was quenched by adding Triton X-100 (2 %). A total of 200 units of DpnII was added to each sample for over-night digestion at 37 °C. dATP, dTTP and dGTP, biotinylated dCTP and 12 µl DNA Polymerase I (Large Klenow fragment) were added before incubation 45 min at 37 °C. A total of 50 unit of T4 DNA ligase along with 7 µL of 20 ng/µL of BSA (Biolabs) and 7 µL of 100 mM ATP were added for 4h at 16°C with constant shaking at 300rpm. After over-night reverse crosslink at 65°C and protein digestion with 5 µL of 10 mg/µl protease K, DNA was extracted by phenol/chloroform purification and ethanol precipitation before resuspension in 100µL of 0.1X TE buffer. Biotin was removed from the unligated fragment using T4 DNA polymerase exonuclease activity. After biotin removal, the samples were purified using AMPure beads with a 1.6 ratio. DNA was fragmented using a Covaris M220 sonicator (peak power 75W, Duty factor 20, Cycles per burst 200, Duration 150 sec). Hi-C libraries were prepared using KAPA LTP Library Preparation Kit (Roche) as in Grob et al. (2014) with 12 amplification cycles. PCR products were purified using AMPure beads (ratio 1.85). Libraries were analyazed using a Qubit (Thermofisher) and a TAPE Station (Agilent) before paired-end 150bp sequencing at the Beijing Genomics Institute (BGI Group; Honk Kong) using a Novaseq Illumina platform. Mapping of Hi-C reads was performed using Hi-C Pro pipeline (Servant et al., 2015) with default pipeline parameters. Data were in visualized using the Juicebox toolsuite (Durand et al., 2016) after SCN normalization (Cournac et al., 2012) and represented in Log10 scale after SCN normalization with Boost-HiC (Carron et al., 2019) setting alpha parameter to 0.2.

RNA-seq

Seedlings grown in long days were fixed in 100% cold acetone under vacuum for 10 min. Cotyledons from 100 plants were dissected and grinded in 2 ml tubes using a Tissue Lyser (Qiagen) for 1 min 30 sec at 30 Hz before RNA extraction using the RNeasy micro kit (Qiagen). RNA was sequenced using the DNBseq platform at the Beijing Genomics Institute (BGI Group) in a 100 bp paired-end mode. For raw data processing, sequencing adaptors were removed from raw reads with trim_galore! v2.10 (https://github.com/FelixKrueger/TrimGalore). Reads were mapped onto combined TAIR10 genome using STAR version 2.7.3a (Dobin et al., 2012) with the following parameters "--alignIntronMin 20 --alignIntronMax 100000 --outFilterMultimapNmax 20 --outMultimapperOrder Random --outFilterMismatchNmax 8 --outSAMtype BAM SortedByCoordinate --outSAMmultNmax 1 --alignMatesGapMax 100000". Gene raw counts were scored using the htseq-count tool from the HTSeq suite version 0.11.3 (Anders et al., 2015) and analyzed with the DESeq2 package (Love et al., 2014) to calculate Log_2 fold change and identify differential expressed genes (p-value < 0.01). TPM (Transcripts per Million) were retrieved by dividing the counts over each gene by its length and the total counts in the sample and multiplying for 10^6. Mean TPM values between two biological replicates were used for subsequent analyses. To draw metagene plots, genes were grouped in expressed or not and expressed genes split in four quantiles of expression with the function ntile() of the R package dplyr (https://CRAN.R-project.org/package=dplyr).

H1 and H3 ChIP-seq experiments

H1.2-GFP and parallel H3 profiling were conducted as in (Fiorucci et al., 2019) with slight modifications to sonicate chromatin to reach mono/di-nucleosome fragment sizes. WT Col-0 or *pH1.2::H1.2-GFP* seedlings were crosslinked for 15 min using 1 % formaldehyde. After dissection, 400 cotyledons were grinded in 2 ml tubes using a Tissue Lyser (Qiagen) for 2 x 1 min at 30 Hz. After resuspension in 100 µL Nuclei Lysis Buffer 0.1 %SDS, the samples were flash frozen in liquid nitrogen and chromatin was sheared using a S220 Focused-ultrasonicator (Covaris) for 17 min at peak power 105 W, Duty factor 5%, 200 cycles per burst, to get fragment sizes between 75 and 300 bp. Immunoprecipitation was performed on 150 µg of chromatin quantified using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific) with 60 µL of Protein-A/G Dynabeads and 3.5 µL of anti-GFP (Thermo Fisher #A11122) for H1.2-GFP and mock (WT) sample or anti-H3 (Abcam #Ab1791) for H3 IPs. Immunoprecipitated DNA was used

for library preparation using the TruSeq® ChIP Sample Preparation Kit (Illumina) and sequenced using a NextSeq 500 system in a single-end 50 bp mode (Genewiz, USA).

H3K27me3 ChIP-Rx

ChIP-Rx analysis of H3K27me3 (Millipore, #07-449) was conducted as in (Nassrallah et al., 2018) using two biological replicates of 8-day-old WT and *2h1* seedlings. For each biological replicate, two independent IPs were carried out using 120 µg of *Arabidopsis* chromatin mixed with 3 % of Drosophila chromatin quantified using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific). DNA samples eluted and purified from the two technical replicates were pooled before library preparation (Illumina TruSeq ChIP) and sequencing (Illumin NextSeq 500, 1x50bp) of all input and IP samples by Fasteris (Geneva, Switzerland).

ChIP analyses

For H3K27me3 spike-in normalized ChIP-Rx, raw reads were pre-processed with Trimmomatic v0.36 (Bolger et al., 2014) to remove Illumina sequencing adapters. 5' and 3' ends with a quality score below 5 (Phred+33) were trimmed and reads shorter than 20 bp after trimming were discarded (trimmomatic-0.36.jar SE -phred33 INPUT.fastq TRIMMED_OUTPUT.fastq ILLUMINACLIP:TruSeq3-SE.fa:2:30:10 LEADING:5 TRAILING:5 MINLEN:20). We aligned the trimmed reads against combined TAIR10 *Arabidopsis thaliana* and *Drosophila melanogaster* (dm6) genomes with Bowtie2v.2.3.2 using "--very-sensitive" setting. Alignments with MAPQ < 10, duplicated reads and reads mapping on repetitive regions (as defined in (Quadrana et al., 2016)) were discarded with sambamba v0.6.8. (Tarasov et al., 2015). Peaks of H3K27me3 read density were called using MACS2 (Zhang et al., 2008) with the command "macs2 callpeak -f BAM --nomodel -q 0.01 -g 120e6 --bw 300 --verbose 3 --broad". Only peaks found in both biological replicates and overlapping for at least 10 % were retained for further analyses. We scored the number of H3K27me3 reads overlapping with marked genes using bedtools v2.29.2 multicov and analyzed them with the DESeq2 package (Love et al., 2014) in the R statistical environment v3.6.2 to identify the genes enriched or depleted in H3K27me3 in mutant mutant plants (p-value < 0.01). To account for differences in sequencing depth we used the function sizeFactors in DESeq2 applying a scaling factor calculated as in (Nassrallah et al., 2018).

For the H1.2-GFP and H3 ChIP-seq datasets, raw reads were processed as for H3K27me3. We counted the reads of GFP-H1.2 and ATAC-seq over genes ant TEs using bedtools v2.29.2 multicov and converted them in median Counts per Million (CPM) dividing the counts over each gene or TE by its length and by the total counts in the sample and multiplying by 10^6. The mean value between biological replicates of IP was used in Figure 1, while the ratio between IP and Input was used for violin-plots analysis of H1.2-GFP in Figure S13. Annotation of genes and TEs overlapping with peaks of histone marks H3K27me3, H3K4me3 and H2Bub were identified using bedtools v2.29.2 intersect. To include nucleosomes in close proximity of the TSS, an upstream region of 250 bp was also considered for the overlap for H3K27me3, TRB1 and H3K4me3. H3K27me3 *TE cluster 1* and *TE cluster 2* were identified using Deeptools plotHeatmap using the --kmeans setting. Tracks were visualized using Integrative Genomics Viewer (IGV) version 2.8.0 (Thorvaldsdóttir et al., 2012). Metagene plots and heatmaps were generated from depth-normalized read densities using Deeptools computeMatrix, plotHeatmap, and plotProfile. Violin-plots, histograms and box-plots were drawn using the package ggplot2 v3.2.1 (https://cran.r-project.org/web/packages/ggplot2/) in R statistical environment. All scripts used will be made publicly available. In Shuffled controls in ChIP analyses, where present, were produced with random permutations of genomic position of the regions if interest. The permutations were generated with bedtools v2.29.2 and the command "bedtools shuffle -chromFirst -seed 28776 -chrom".

For MNase-seq analyses, MNase read density (Choi et al, 2017) was obtained from NCBI GEO under the accession GSE96994. Genomic location of WPNs shared between WT and *2h1* plants were identified as overlapping WPN coordinates between the two genotypes calculated with bedtools v2.29.2 intersect.

Telobox positioning was analyzed using the coordinates described in (Zhou et al., 2018) and obtained from https://gbrowse.mpipz.mpg.de/cgi-bin/gbrowse/arabidopsis10_turck_public/?l=telobox;f=save+datafile. *Telobox* repeat numbers were scored over 10-bp non-overlapping bins, smoothed with a 50-bp sliding window and subsequently used to plot telobox density.

TRB1 ChIP-qPCR

Four biological replicates of 8-day-old WT or *2h1* crosslinked seedlings were ground in liquid nitrogen. For each replicate, 2 g of tissue was resuspended in 30 mL EB1. Nuclei lysis buffer was added to 520 µg of chromatin to reach 1 µg/µL of proteins according to BCA assay. 20 µl was kept as an input. 1.44mL of Dynabeads protein G slurry (ThermoFisher Scientific 10004D) was added to each sample and incubated 1h at 4°C with 6.5 µl of Bridging Antibody for Mouse IgG (Active Motif #53017) on a rotating wheel. 60 µl of bridged beads were then added to each sample for pre-clearing. The remaining bridged beads were incubated with 50µL of anti-TRB1 5.2 antibody (Schrumpfová et al., 2014) for 3 hours at 4°C under rotation before transfer to the pre-cleared chromatin samples and incubation overnight at 4°C under rotation. Beads were washed and chromatin was eluted in 500 µl of SDS elution buffer (1 % SDS, 0.1 M NaHCO3) at 65°C before reverse crosslinking adding 20 µl of 5 M NaCl overnight at 65°C. After Proteinase K digestion at 45°C for 1 h, DNA was purified by phenol-chloroform extraction and ethanol precipitated. The pellet of each input and IP was resuspended in 40 µl of TE pH 8.0. DNA was analyzed by quantitative PCR using a LightCycler 480 SYBR green I Master mix and a LightCycler 480 (Roche) using the primers sequences given in Additional file 3.

ChIP-hybrization analysis of telomeric H3K27me3 and H3

Anti-H3K27me3 (Millipore, #07-449 antibody) and anti-H3 (Abcam #Ab1791 antibody) ChIPs were conducted using 2 g of tissue. Pellets of both inputs (20%) and immunoprecipitated DNA were resuspended in 40 µl of TE, pH 8.0 and analyzed through dot-blot hybridization using radioactively labeled telomeric probe synthesized by non-template PCR (IJdo et al., 1991; Adamusová et al. 2020). ITRs contribution to the hybridization signal was minimized using high stringency hybridization as detailed in Adamusová et al. (2020).

DNA sequence motif search

Motifs enriched in gene promoters (-500 bp to +250 bp after the TSS) and in annotated units of *TE cluster 1* elements were identified using MEME version 5.1.1 (Bailey et al., 2015). The following options were used for promoters: "-dna -mod anr -revcomp -nmotifs 10 -minw 5 -maxw 9" and for TEs: "-dna -mod anr -nmotifs 10 -minw 5 -maxw 9 -objfun de -neg Araport11_AllTEs.fasta -revcomp -markov_order 0 -maxsites 10000" where Araport11_AllTEs.fasta correspond to the fasta sequence of all TEs annotated in Araport11.

Gene ontology analysis

Gene ontology analysis of H3K27me3 differentially marked genes were retrieved using the GO-TermFinder software (Boyle et al., 2004) via the Princeton GO-TermFinder interface (http://go.princeton.edu/cgi-bin/GOTermFinder). The REVIGO (Supek et al., 2011) platform was utilized to reduce the number of GO terms and redundant terms were further manually filtered. The log10 p-values of these unique GO terms were then plotted with pheatmap (https://CRAN.R-project.org/package=pheatmap) with no clustering.

Protein alignment

Protein sequences of H1.1, H1.2, H1.3, TRB1, TRB2 and TRB3 were aligned using T-Coffee (http://tcoffee.crg.cat/apps/tcoffee/do:regular) with default parameters. Pairwise comparison for similarity and identy score were calculated using Ident and Sim tool (https://www.bioinformatics.org/sms2/ident_sim.html).

Genomic data

All public genomic data are listed in Additional file 4. Additional files will be made available upon request to the corresponding authors.

Funding

FB benefitted from grants of the Agence Nationale de la Recherche (France) (projects ANR-10-LABX-54, ANR-18-CE13-0004-01, ANR-17-CE12-0026-02). Collaborative work between FB and CeB was supported by a research grant from the Velux Foundation (Switzerland). Collaborative work between FB and AP was supported by CNRS EPIPLANT Action (France). GT benefitted from a short-term fellowship of the COST Action CA16212 INDEPTH (EU) for training in Hi-C by SG in UG's laboratory, which is supported by the University of Zurich (Switzerland) and the Swiss National Science Foundation (project 31003A_179553). SA benefitted from a CAP20-25 Emergence research grant from Région Auvergne-Rhône-Alpes (France). Work in JF team was supported by the Czech Science Foundation (project 20-01331X) and Ministry of Education, Youth and Sports of the Czech Republic - project INTER-COST (LTC20003).

Author contributions

GT, LW and ClB performed ChIP experiments; GT and MB generated ATAC-seq datasets; KA and MF performed telomere dot-blots; SA performed cytological experiments and quantification; GT and SG generated the Hi-C datasets. AK and VC developed ATAC-seq bioinformatics tools. LoC, GT, ClB performed RNA-seq, ChIP-seq, and ATAC-seq bioinformatics analyses. LoC and LeC performed Hi-C bioinformatics. GT, ClB, SG, and FB conceived the study. FB, ClB, AC, CeB, ChB, AC, SA, AP, VC, UG, PPS, JF, and SG supervised research, discussed the results and edited the manuscript. FB, ClB and GT wrote the paper.

Acknowledgements

The authors are very grateful to Erwann Cailleux (IBENS, Paris, France) and David Latrasse (IPS2, Orsay, France) for technical guidance with ATAC-seq; to Magali Charvin (IBENS, Paris, France) for technical assistance with the IBENS plant growth facility, to Frédérique Perronet (IBPS, France) for providing Drosophila samples; Kinga Rutowicz (University of Zurich, Switzerland) and Angélique Déléris (IBENS, Paris and I2BC, Gif-Sur-Yvette, France) for sharing unpublished work.

References

Achrem, M., Szućko, I., and Kalinka, A. (2020). The epigenetic regulation of centromeres and telomeres in plants and animals. Comparative Cytogenetics *14*, 265–311.

Adamusová, K., Khosravi, S., Fujimoto, S., Houben, A., Matsunaga, S., Fajkus, J., and Fojtová, M. (2020). Two combinatorial patterns of telomere histone marks in plants with canonical and non-canonical telomere repeats. Plant Journal *102*, 678–687.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Ascenzi, R., and Gantt, J.S. (1999). Subnuclear distribution of the entire complement of linker histone variants in Arabidopsis thaliana. Chromosoma *108*, 345–355.

Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME Suite. Nucleic Acids Res *43*, W39–W49.

Bednar, J., Hamiche, A., and Dimitrov, S. (2016). H1-nucleosome interactions and their functional implications. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms *1859*, 436–443.

Bednar, J., Garcia-Saez, I., Boopathi, R., Cutter, A.R., Papai, G., Reymer, A., Syed, S.H., Lone, I.N., Tonchev, O., Crucifix, C., et al. (2017). Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. Molecular Cell *66*, 384-397.e8.

Beh, L.Y., Colwell, L.J., and Francis, N.J. (2012). A core subunit of Polycomb repressive complex 1 is broadly conserved in function but not primary sequence. Proceedings of the National Academy of Sciences of the United States of America *109*.

Bernatavichute, Y.V., Zhang, X.Y., Cokus, S., Pellegrini, M., and Jacobsen, S.E. (2008). Genome-wide association of histone H3 lysine nine methylation with CHG DNA methylation in Arabidopsis thaliana. PLoS One *8*.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nature Reviews. Genetics *17*, 661–678.

Boyle, E.I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J.M., and Sherlock, G. (2004). GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Bioinformatics *20*, 3710–3715.

Cao, K., Lailler, N., Zhang, Y., Kumar, A., Uppal, K., Liu, Z., Lee, E.K., Wu, H., Medrzycki, M., Pan, C., et al. (2013). High-Resolution Mapping of H1 Linker Histone Variants in Embryonic Stem Cells. PLOS Genetics *9*, e1003417.

Carron, L., Morlot, J.B., Matthys, V., Lesne, A., Mozziconacci, J., and Birol, I. (2019). Boost-HiC: Computational enhancement of long-range contacts in chromosomal contact maps. Bioinformatics *35*, 2724–2729.

Catez, F., Yang, H., Tracey, K.J., Reeves, R., Misteli, T., and Bustin, M. (2004). Network of Dynamic Interactions between Histone H1 and High-Mobility-Group Proteins in Chromatin. Molecular and Cellular Biology *24*, 4321–4328.

Charbonnel, C., Rymarenko, O., Da Ines, O., Benyahya, F., White, C.I., Butter, F., and Amiard, S. (2018). The Linker Histone GH1-HMGA1 Is Involved in Telomere Stability and DNA Damage Repair. Plant Physiology *177*, 311–327.

Choi, J., Lyons, D.B., Kim, M.Y., Moore, J.D., and Zilberman, D. (2019). DNA Methylation and Histone H1 Jointly Repress Transposable Elements and Aberrant Intragenic Transcripts. Molecular Cell.

Copenhaver, G.P., and Pikaard, C.S. (1996). RFLP and physical mapping with an rDNA-specific endonuclease reveals that nucleolus organizer regions of Arabidopsis thaliana adjoin the telomeres on chromosomes 2 and 4. The Plant Journal *9*, 259–272.

Cournac, A., Marie-Nelly, H., Marbouty, M., Koszul, R., and Mozziconacci, J. (2012). Normalization of a chromosomal contact map. BMC Genomics *13*, 436.

Déjardin, J., and Kingston, R.E. (2009). Purification of Proteins Associated with Specific Genomic Loci. Cell *136*, 175–186.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2012). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. Cell Syst *3*, 99–101.

Dvořáčková, M., Rossignol, P., Shaw, P.J., Koroleva, O.A., Doonan, J.H., and Fajkus, J. (2010). AtTRB1, a telomeric DNA-binding protein from Arabidopsis, is concentrated in the nucleolus and shows highly dynamic association with chromatin. Plant Journal *61*, 637–649.

Dvořáčková, M., Fojtová, M., and Fajkus, J. (2015). Chromatin dynamics of plant telomeres and ribosomal genes. Plant Journal *83*, 18–37.

Fajkus, J., and Trifonov, E.N. (2001). Columnar packing of telomeric nucleosomes. Biochemical and Biophysical Research Communications *280*, 961–963.

Fajkus, J., Kovařík, A., MKrálovics, R., and Bezděk, M. (1995). Organization of telomeric and subtelomeric chromatin in the higher plant Nicotiana tabacum. MGG Molecular & General Genetics *247*, 633–638.

Feng, S., Cokus, S.J., Schubert, V., Zhai, J., Pellegrini, M., and Jacobsen, S.E. (2014). Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in Arabidopsis. Molecular Cell *55*, 694–707.

Fiorucci, A.-S., Bourbousse, C., Concia, L., Rougée, M., Deton-Cabanillas, A.-F., Zabulon, G., Layat, E., Latrasse, D., Kim, S.K., Chaumont, N., et al. (2019). Arabidopsis S2Lb links AtCOMPASS-like and SDG2 activity in H3K4me3 independently from histone H2B monoubiquitination. Genome Biology *20*, 100.

Fitzgerald, M.S., Riha, K., Gao, F., Ren, S., McKnight, T.D., and Shippen, D.E. (1999). Disruption of the telomerase catalytic subunit gene from Arabidopsis inactivates telomerase and leads to a slow loss of telomeric DNA. Proc Natl Acad Sci U S A *96*, 14813–14818.

Fojtová, M., and Fajkus, J. (2014). Epigenetic regulation of telomere maintenance. Cytogenetic and Genome Research *143*, 125–135.

Francis, N.J., Kingston, R.E., and Woodcock, C.L. (2004). Chromatin compaction by a polycomb group protein complex. Science *306*, 1574–1577.

Fransz, P., de Jong, J.H., Lysak, M., Castiglione, M.R., and Schubert, I. (2002). Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate 10.1073/pnas.212325299. Proceedings of the National Academy of Sciences *99*, 14584–14589.

Fyodorov, D.V., Zhou, B.-R., Skoultchi, A.I., and Bai, Y. (2017). Emerging roles of linker histones in regulating chromatin structure and function. Nature Reviews Molecular Cell Biology *19*, 192.

Galati, A., Micheli, E., and Cacchione, S. (2013). Chromatin structure in telomere dynamics. Frontiers in Oncology *3 MAR*, 1–16.

Geeven, G., Zhu, Y., Kim, B.J., Bartholdy, B.A., Yang, S.M., Macfarlan, T.S., Gifford, W.D., Pfaff, S.L., Verstegen, M.J.A.M., Pinto, H., et al. (2015). Local compartment changes and regulatory landscape alterations in histone H1-depleted cells. Genome Biology *16*.

Grafi, G., Ben-Meir, H., Avivi, Y., Moshe, M., Dahan, Y., and Zemach, A. (2007). Histone methylation controls telomerase-independent telomere lengthening in cells undergoing dedifferentiation. Developmental Biology *306*, 838–846.

Grau, D.J., Chapman, B.A., Garlick, J.D., Borowsky, M., Francis, N.J., and Kingston, R.E. (2011). Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. Genes & Development *25*, 2210–2221.

Grob, S., Schmid, M.W., and Grossniklaus, U. (2014). Hi-C analysis in Arabidopsis identifies the KNOT, a structure with similarities to the flamenco locus of Drosophila. Molecular Cell *55*, 678–693.

Grossniklaus, U., and Paro, R. (2014). Transcriptional silencing by polycomb-group proteins. Cold Spring Harb. Perspect. Biol. 6, a019331.

He, S., Vickers, M., Zhang, J., and Feng, X. (2019). Natural depletion of H1 in sex cells causes DNA demethylation, heterochromatin decondensation and transposon activation. ELife *8*, 1–23.

Hergeth, S.P., and Schneider, R. (2015). The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. EMBO Reports *16*, 1439–1453.

Hsieh, P.H., He, S., Buttress, T., Gao, H., Couchman, M., Fischer, R.L., Zilberman, D., and Feng, X. (2016). Arabidopsis male sexual lineage exhibits more robust maintenance of CG methylation than somatic tissues. Proceedings of the National Academy of Sciences of the United States of America *113*, 15132–15137.

Hugues, A., Jacobs, C.S., and Roudier, F. (2020). Mitotic Inheritance of PRC2-Mediated Silencing: Mechanistic Insights and Developmental Perspectives. Frontiers in Plant Science *11*, 1–11.

IJdo, J.W., Wells, R.A., Baldini, A., and Reeders, S.T. (1991). Improved telomere detection using a telomere repeat probe (TTAGGG) n generated by PCR. Nucleic Acids Research *19*, 4780–4780.

Illingworth, R.S. (2019). Chromatin folding and nuclear architecture: PRC1 function in 3D.

Izzo, A., Kamieniarz-Gdula, K., Ramírez, F., Noureen, N., Kind, J., Manke, T., van Steensel, B., and Schneider, R. (2013). The Genomic Landscape of the Somatic Linker Histone Subtypes H1.1 to H1.5 in Human Cells. Cell Reports *3*, 2142–2154.

Jamieson, K., Mcnaught, K.J., Ormsby, T., Leggett, N.A., Honda, S., and Selker, E.U. (2018). Telomere repeats induce domains of H3K27 methylation in Neurospora. ELife *7*, 1–18.

Kim, J.M., Kim, K., Punj, V., Liang, G., Ulmer, T.S., Lu, W., and An, W. (2015). Linker histone H1.2 establishes chromatin compaction and gene silencing through recognition of H3K27me3. Scientific Reports *5*, 1–16.

Kotliński, M., Knizewski, L., Muszewska, A., Rutowicz, K., Lirski, M., Schmidt, A., Baroux, C., Ginalski, K., and Jerzmanowski, A. (2017). Phylogeny-based systematization of arabidopsis proteins with histone H1 globular domain. Plant Physiology *174*, 27–34.

Krishnakumar, R., Gamble, M.J., Frizzell, K.M., Berrocal, J.G., Kininis, M., and Kraus, W.L. (2008). Reciprocal binding of PARP-1 and histone H1 at promoters specifies transcriptional outcomes. Science *319*, 819–821.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods *9*, 357–359.

Lau, M.S., Schwartz, M.G., Kundu, S., Savol, A.J., Wang, P.I., Marr, S.K., Grau, D.J., Schorderet, P., Sadreyev, R.I., Tabin, C.J., et al. (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. Science *355*, 1081.

Lejnine, S., Makarov, V.L., and Langmore, J.P. (1995). Conserved nucleoprotein structure at the ends of vertebrate and invertebrate chromosomes. Proceedings of the National Academy of Sciences of the United States of America *92*, 2393–2397.

Liu, C., Wang, C., Wang, G., Becker, C., Zaidem, M., and Weigel, D. (2016). Genome-wide analysis of chromatin packing in Arabidopsis thaliana at single-gene resolution. Genome Research.

Liu, S., de Jonge, J., Trejo-Arellano, M.S., Santos-González, J., Köhler, C., and Hennig, L. (2020). Role of H1 and DNA methylation in selective regulation of transposable elements during heat stress. New Phytologist nph.17018.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*, 550.

Lu, Z., Hofmeister, B.T., Vollmers, C., Dubois, R.M., Schmitz, R.J., and Schmitz, J. (2016). Combining ATAC-seq with nuclei sorting for discovery of cis-regulatory regions in plant genomes. Nucleic Acids Research *45*, 1–13.

de Lucas, M., Pu, L., Turco, G., Gaudinier, A., Morao, A.K., Harashima, H., Kim, D., Ron, M., Sugimoto, K., Roudier, F., et al. (2016). Transcriptional Regulation of Arabidopsis Polycomb Repressive Complex 2 Coordinates Cell-Type Proliferation and Differentiation. Plant Cell *28*, 2616.

Lyons, D.B., and Zilberman, D. (2017). DDM1 and lsh remodelers allow methylation of DNA wrapped in nucleosomes. ELife *6*, 1–20.

Ma, Z., Castillo-González, C., Wang, Z., Sun, D., Hu, X., Shen, X., Potok, M.E., and Zhang, X. (2018). Arabidopsis Serrate Coordinates Histone Methyltransferases ATXR5/6 and RNA Processing Factor RDR6 to Regulate Transposon Expression. Developmental Cell *45*, 769–784.e6.

Majerová, E., Mandáková, T., Vu, G.T.H., Fajkus, J., Lysak, M.A., and Fojtová, M. (2014). Chromatin features of plant telomeric sequences at terminal vs. internal positions. Frontiers in Plant Science *5*, 1–10.

Makarov, V.L., Lejnine, S., Bedoyan, J., and Langmore, J.P. (1993). Nucleosomal organization of telomere-specific chromatin in rat. Cell *73*, 775–787.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. Molecular Cell *32*, 503–518.

Martin, C., Cao, R., and Zhang, Y. (2006). Substrate preferences of the EZH2 histone methyltransferase complex. Journal of Biological Chemistry *281*, 8365–8370.

Moissiard, G., Cokus, S.J., Cary, J., Feng, S., Billi, A.C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B.R., McCord, R.P., et al. (2012). MORC Family ATPases Required for Heterochromatin Condensation and Gene Silencing. Science.

Mozgová, I., Schrumpfová, P.P., Hofr, C., and Fajkus, J. (2008). Functional characterization of domains in AtTRB1, a putative telomere-binding protein in Arabidopsis thaliana. Phytochemistry *69*, 1814–1819.

Nassrallah, A., Rougée, M., Bourbousse, C., Drevensek, S., Fonseca, S., Iniesto, E., Ait-Mohamed, O., Deton-Cabanillas, A.F., Zabulon, G., Ahmed, I., et al. (2018). DET1-mediated degradation of a SAGA-like deubiquitination module controls H2Bub homeostasis. ELife *7*, 1–29.

Over, R.S., and Michaels, S.D. (2014). Open and closed: The roles of linker histones in plants and animals. Molecular Plant *7*, 481–491.

Papareddy, R.K., Páldi, K., Paulraj, S., Kao, P., Lutzmayer, S., and Nodine, M.D. (2020). Chromatin regulates expression of small RNAs to help maintain transposon methylome homeostasis in Arabidopsis. Genome Biology *21*, 251.

Pontvianne, F., and Grob, S. (2020). Three-dimensional nuclear organization in Arabidopsis thaliana. Journal of Plant Research.

Probst, A.V., Desvoyes, B., and Gutierrez, C. (2020). Similar yet critically different: the distribution, dynamics and function of histone variants. Journal of Experimental Botany *71*, 5191–5204.

Quadrana, L., Bortolini Silveira, A., Mayhew, G.F., LeBlanc, C., Martienssen, R.A., Jeddeloh, J.A., and Colot, V. (2016). The Arabidopsis thaliana mobilome and its impact at the species level. ELife *5*, e15716.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res *44*, W160–W165.

Regad, F., Lebas, M., and Lescure, B. (1994). Interstitial Telomeric Repeats within the Arabidopsis thaliana Genome. Journal of Molecular Biology *239*, 163–169.

Richards, E.J., and Ausubel, F.M. (1988). Isolation of a higher eukaryotic telomere from Arabidopsis thaliana. Cell *53*, 127–136.

Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., Bouyer, D., Caillieux, E., Duvernois-Berthet, E., and Al-Shikhley, L. (2011). Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. The EMBO Journal *30*, 1928–1938.

Rutowicz, K., Puzio, M., Halibart-Puzio, J., Lirski, M., Kroteń, M.A., Kotliński, M., Kniżewski, Ł., Lange, B., Muszewska, A., Śniegowska-Świerk, K., et al. (2015). A specialized histone H1 variant is required for adaptive responses to complex abiotic stress and related DNA methylation in Arabidopsis.

Rutowicz, K., Lirski, M., Mermaz, B., Teano, G., Schubert, J., Mestiri, I., Kroteń, M.A., Fabrice, T.N., Fritz, S., Grob, S., et al. (2019). Linker histones are fine-scale chromatin architects modulating developmental decisions in Arabidopsis. Genome Biology *20*, 157.

Santos, A.P., Gaudin, V., Mozgová, I., Pontvianne, F., Schubert, D., Tek, A.L., Dvořáčková, M., Liu, C., Fransz, P., Rosa, S., et al. (2020). Tiding-up the plant nuclear space: domains, function and dynamics. Journal of Experimental Botany 1–19.

Schep, A.N., Buenrostro, J.D., Denny, S.K., Schwartz, K., Sherlock, G., and Greenleaf, W.J. (2015). Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. Genome Research *25*, 1757–1770.

Schrumpfová, P., Kuchař, M., Miková, G., Skříšovská, L., Kubičárová, T., and Fajkus, J. (2004). Characterization of two Arabidopsis thaliana myb-like proteins showing affinity to telomeric DNA sequence. Genome *47*, 316–324.

Schrumpfová, P.P., Vychodilová, I., Dvořáčková, M., Majerská, J., Dokládal, L., Schořová, Š., and Fajkus, J. (2014). Telomere repeat binding proteins are functional components of Arabidopsis telomeres and interact with telomerase. Plant Journal *77*, 770–781.

Schrumpfová, P.P., Vychodilová, I., Hapala, J., Schořová, Š., Dvořáček, V., and Fajkus, J. (2016). Telomere binding protein TRB1 is associated with promoters of translation machinery genes in vivo. Plant Molecular Biology *90*, 189–206.

Schrumpfová, P.P., Fojtová, M., and Fajkus, J. (2019). Telomeres in Plants and Humans: Not So Different, Not So Similar. Cells *8*, 58.

Schubert, D. (2019). Evolution of Polycomb-group function in the green lineage. F1000Res *8*.

Schuettengruber, B., Bourbon, H.M., Di Croce, L., and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell *171*, 34–57.

Sequeira-Mendes, J., Araguez, I., Peiro, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S.E., Bastolla, U., and Gutierrez, C. (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. The Plant Cell *26*, 2351–2366.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: An optimized and flexible pipeline for Hi-C data processing. Genome Biology *16*, 1–11.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., and Hoichman, M. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C., Bender, W., and Kingston, R.E. (1999). Stabilization of Chromatin Structure by PRC1, a Polycomb Complex. Cell *98*, 37–46.

She, W., and Baroux, C. (2015). Chromatin dynamics in pollen mother cells underpin a common scenario at the somatic-to-reproductive fate transition of both the male and female lineages in Arabidopsis. Frontiers in Plant Science *6*, 294.

She, W., Grimanelli, D., Rutowicz, K., Whitehead, M.W.J., Puzio, M., Kotliński, M., Jerzmanowski, A., and Baroux, C. (2013). Chromatin reprogramming during the somatic-to-reproductive cell fate transition in plants. Development (Cambridge) *140*, 4008–4019.

Shu, H., Wildhaber, T., Siretskiy, A., Gruissem, W., and Hennig, L. (2012). Distinct modes of DNA accessibility in plant chromatin. Nature Communications *3*, 1281.

Simpson, R.T. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry *17*, 5524–5531.

Sun, L., Jing, Y., Liu, X., Li, Q., Xue, Z., Cheng, Z., Wang, D., He, H., and Qian, W. (2020). Heat stress-induced transposon activation correlates with 3D chromatin organization rearrangement in Arabidopsis. Nature Communications *11*, 1886.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLOS ONE *6*, e21800.

Tan, L., Zhang, C., Hou, X., Shao, C., Lu, Y., Zhou, J., Li, Y., Li, L., Chen, S., and He, X. (2018). The PEAT protein complexes are required for histone deacetylation and heterochromatin silencing. The EMBO Journal *37*, 1–21.

Tarasov, A., Vilella, A.J., Cuppen, E., Nijman, I.J., and Prins, P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics *31*, 2032–2034.

Terranova, R., Yokobayashi, S., Stadler, M.B., Otte, A.P., van Lohuizen, M., Orkin, S.H., and Peters, A.H.F.M. (2008). Polycomb Group Proteins Ezh2 and Rnf2 Direct Genomic Contraction and Imprinted Repression in Early Mouse Embryos. Developmental Cell *15*, 668–679.

Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2012). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics *14*, 178–192.

Tremousaygue, D., Manevski, A., Bardet, C., Lescure, N., and Lescure, B. (1999). Plant interstitial telomere motifs participate in the control of gene expression in root meristems. Plant Journal *20*, 553–561.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007). Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLoS Genetics *3*, e86.

Uchida, W., Matsunaga, S., Sugiyama, R., and Kawano, S. (2002). Interstitial telomere-like repeats in the Arabidopsis thaliana genome. Genes and Genetic Systems *77*, 63–67.

Vannier, J.B., Depeiges, A., White, C., and Gallego, M.E. (2009). ERCC1/XPF protects short telomeres from homologous recombination in Arabidopsis thaliana. PLoS Genetics *5*, 1–11.

Vaquero-Sedas, M.I., Gámez-Arjona, F.M., and Vega-Palas, M.A. (2011). Arabidopsis thaliana telomeres exhibit euchromatic features. Nucleic Acids Research *39*, 2007–2017.

Vaquero-Sedas, M.I., Luo, C., and Vega-Palas, M.A. (2012). Analysis of the epigenetic status of telomeres by using ChIP-seq data. Nucleic Acids Research *40*.

Vega-Vaquero, A., Bonora, G., Morselli, M., Vaquero-Sedas, M.I., Rubbi, L., Pellegrini, M., and Vega-Palas, M.A. (2016). Novel features of telomere biology revealed by the absence of telomeric DNA methylation. Genome Research *26*, 1047–1056.

Veluchamy, A., Jégu, T., Ariel, F., Latrasse, D., Mariappan, K.G., Kim, S.-K., Crespi, M., Hirt, H., Bergounioux, C., Raynaud, C., et al. (2016). LHP1 Regulates H3K27me3 Spreading and Shapes the Three-Dimensional Conformation of the Arabidopsis Genome. PLOS ONE *11*, 1–25.

Wang, C., Liu, C., Roqueiro, D., Grimm, D., Schwab, R., Becker, C., Lanz, C., and Weigel, D. (2014). Genome-wide analysis of local chromatin packing in Arabidopsis thaliana. Genome Research.

Wang, C., Liu, C., Roqueiro, D., Grimm, D., Schwab, R., Becker, C., Lanz, C., and Weigel, D. (2015). Genome-wide analysis of local chromatin packing in Arabidopsis thaliana. Genome Research *25*, 246–256.

Wollmann, H., Stroud, H., Yelagandula, R., Tarutani, Y., Jiang, D., Jing, L., Jamge, B., Takeuchi, H., Holec, S., Nie, X., et al. (2017). The histone H3 variant H3.3 regulates gene body DNA methylation in Arabidopsis thaliana. Genome Biology *18*, 1–10.

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J.D., Pai, A., Song, C., Zhuang, M., Klasfeld, S., He, C., et al. (2017). Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. Nature Genetics *49*, 1546–1552.

Yan, W., Chen, D., Smaczniak, C., Engelhorn, J., Liu, H., Yang, W., Graf, A., Carles, C.C., Zhou, D.X., and Kaufmann, K. (2018). Dynamic and spatial restriction of Polycomb activity by plant histone demethylases. Nature Plants *4*, 681–689.

Yuan, W., Wu, T., Fu, H., Dai, C., Wu, H., Liu, N., Li, X., Xu, M., Zhang, Z., Niu, T., et al. (2012). Dense chromatin activates polycomb repressive complex 2 to regulate H3 lysine 27 methylation. Science *337*, 971–975.

Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L., and Zilberman, D. (2013). The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell *153*, 193–205.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y.V., Pellegrini, M., Goodrich, J., and Jacobsen, S.E. (2007). Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in Arabidopsis. PLOS Biology *5*, e129.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology *9*, R137.

Zhou, B.-R., Jiang, J., Feng, H., Ghirlando, R., Xiao, T.S., and Bai, Y. (2015). Structural Mechanisms of Nucleosome Recognition by Linker Histones. Molecular Cell *59*, 628–638.

Zhou, Y., Hartwig, B., James, G.V., Schneeberger, K., and Turck, F. (2016). Complementary activities of TELOMERE REPEAT BINDING proteins and polycomb group complexes in transcriptional regulation of target genes. Plant Cell *28*, 87–101.

Zhou, Y., Wang, Y., Krause, K., Yang, T., Dongus, J.A., Zhang, Y., and Turck, F. (2018). Telobox motifs recruit CLF/SWN-PRC2 for H3K27me3 deposition via TRB factors in Arabidopsis. Nature Genetics *50*, 638–644.

Figure S1

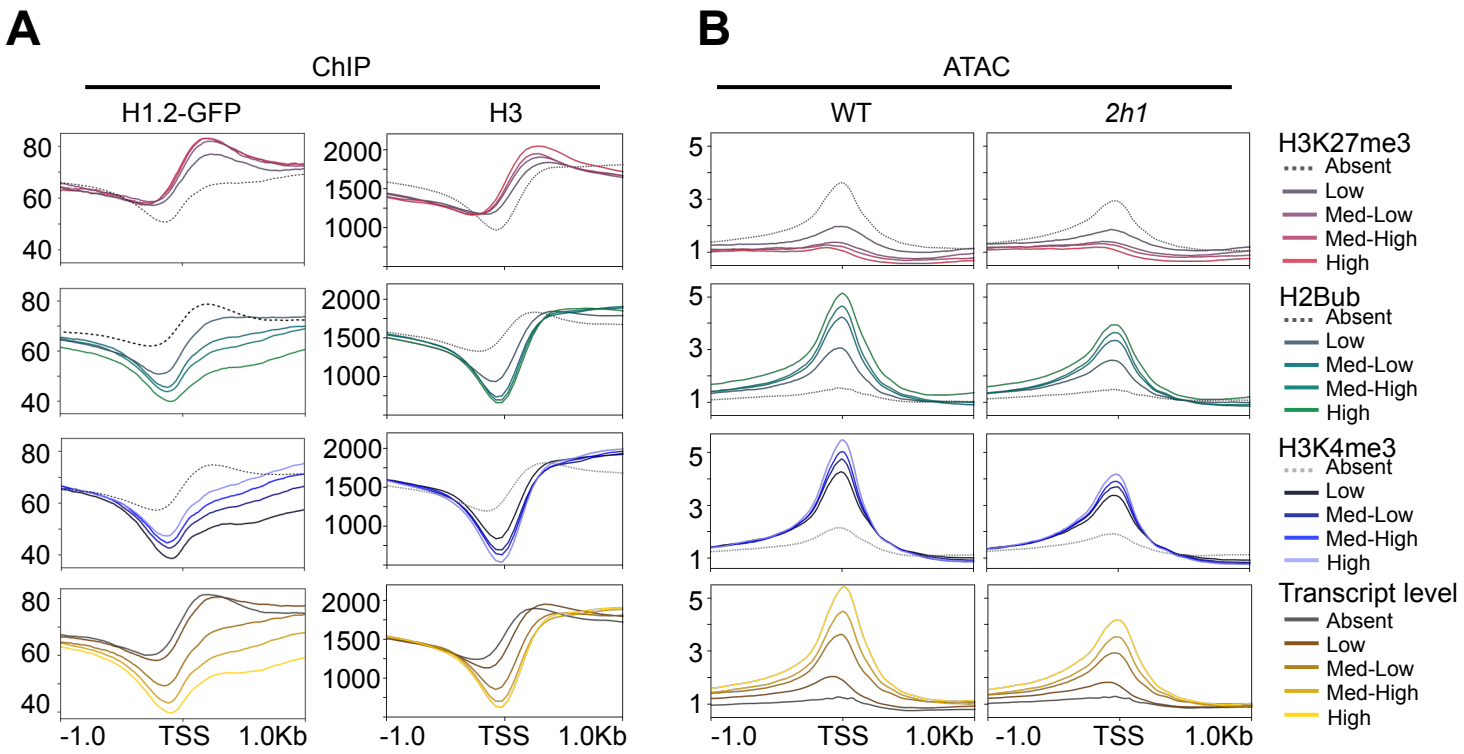


Figure S1. H1.2 distribution and DNA accessibility on distinct chromatin states.

A. H1.2-GFP and H3 enrichment (mean IP/Input signal) over genes marked by different histone modifications characteristic of PRC2-based repression (H3K27me3; n=7542), of transcription initiation (H3K4me3; n=18735), transcription elongation (H2Bub; n=11357) or according to gene expression. Genes with no detectable reads in our RNA-seq analyses of WT plants were considered as not expressed genes (n=5894) as compared to other genes (n=22103). Data represent the mean value of two biological replicates. All ChIPs have been generated in this study except H2Bub and H3K4me3 (Additional file 4). **B.** Same analysis as in (A) for ATAC-seq mean read coverage (reads per million) used as a proxy of nucleosome accessibility in WT and *2h1* plants.

Figure S2

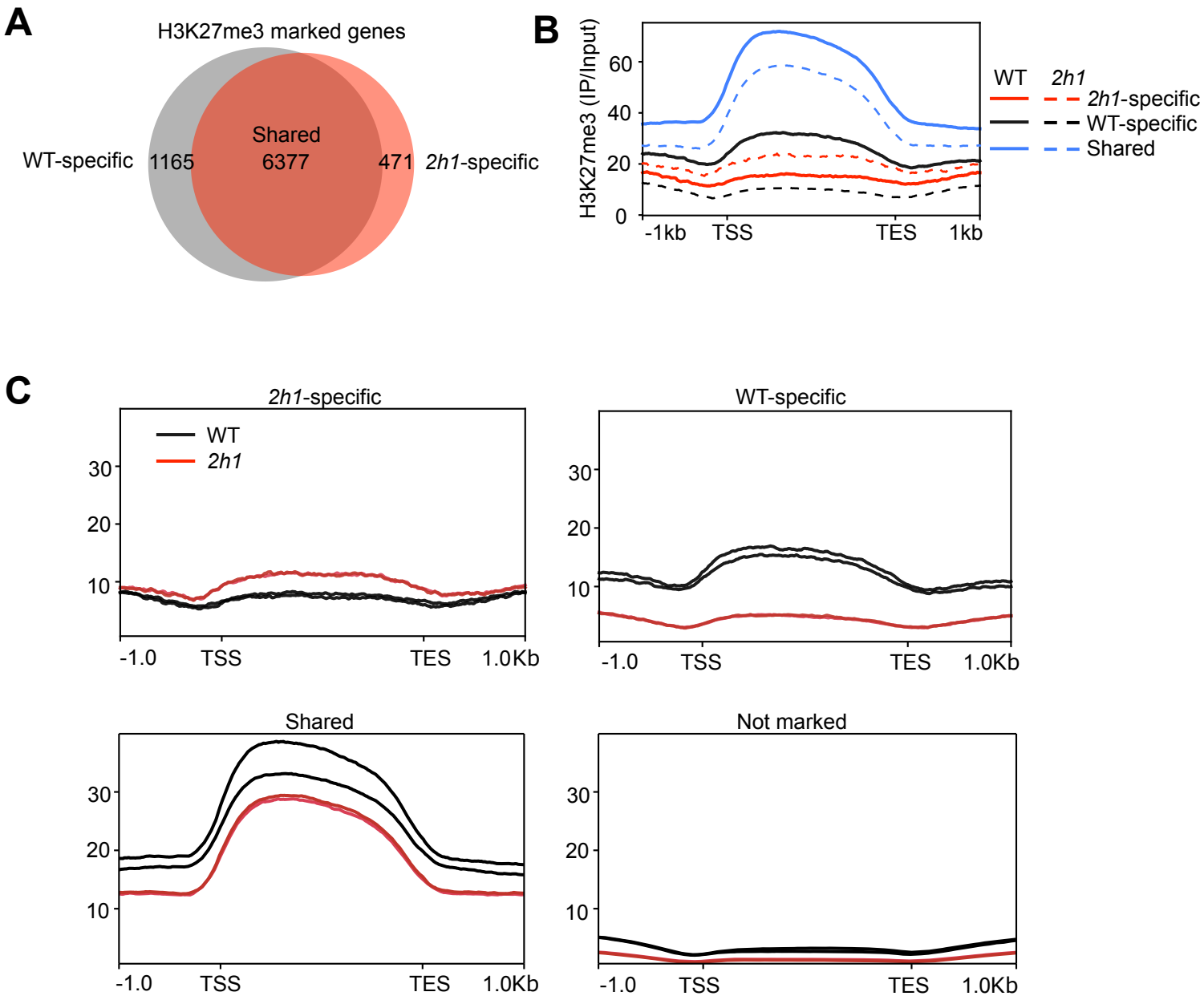


Figure S2. Comparison of the gene repertoire significantly marked by H3K27me3 in WT and *2h1* plants.
A. Number of H3K27me3-marked genes in WT and *2h1* seedlings. **B.** H3K27me3 profiles over the sets of genes acquiring (*2h1*-specific) or losing (WT-specific) H3K27me3 marking in *2h1* plants. Gene sets are defined as in (A). **C.** Same analysis as in (B) showing consistency between the two independent biological replicates.

Figure S3

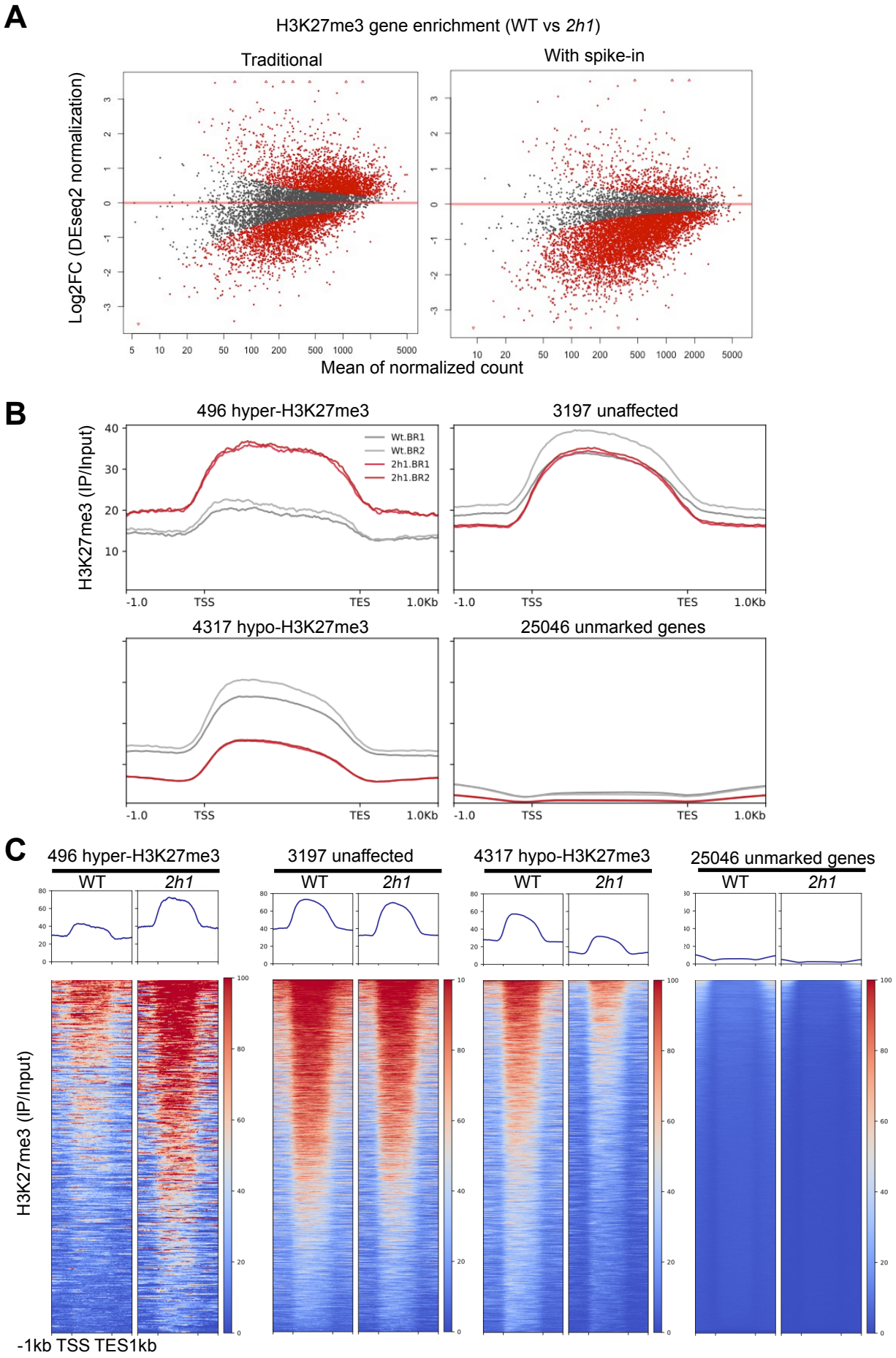


Figure S3. Detailed analysis of H3K27me3 ChIP-Rx differential analyses. A. Comparison of the DEseq2 results of two independent biological replicates (BR) using either a spike-in normalization factor or traditional Deseq2 normalization. **B.** Variability of H3K27me3 profiles on differentially marked genes among the two independent biological replicates. **C.** Detailed analysis of data presented in Figure 2D. In each cluster genes were ranked according to mean H3K27me3 level.

Figure S4

1.9e-051 1094 motifs in 418 promoters (84%)

3.2e-004 114 motifs in 83 promoters (17%)

1.8e-006 104 motifs in 87 promoters (17%)

Figure S4. Sequence motifs over-represented in the 496 H3K27me3 genes hyper-marked by H3K27me3 in *2h1* plants (E-value < 1e-220). E-values were calculated against random sequences. The two first motifs could not be matched to any previously known regulatory motif while the 3rd identified motif corresponds to the previously described *telobox* motif (AAACCCTA). In total, 104 *teloboxes* were found distributed within 87 of the 496 promoters (17%).

Figure S5

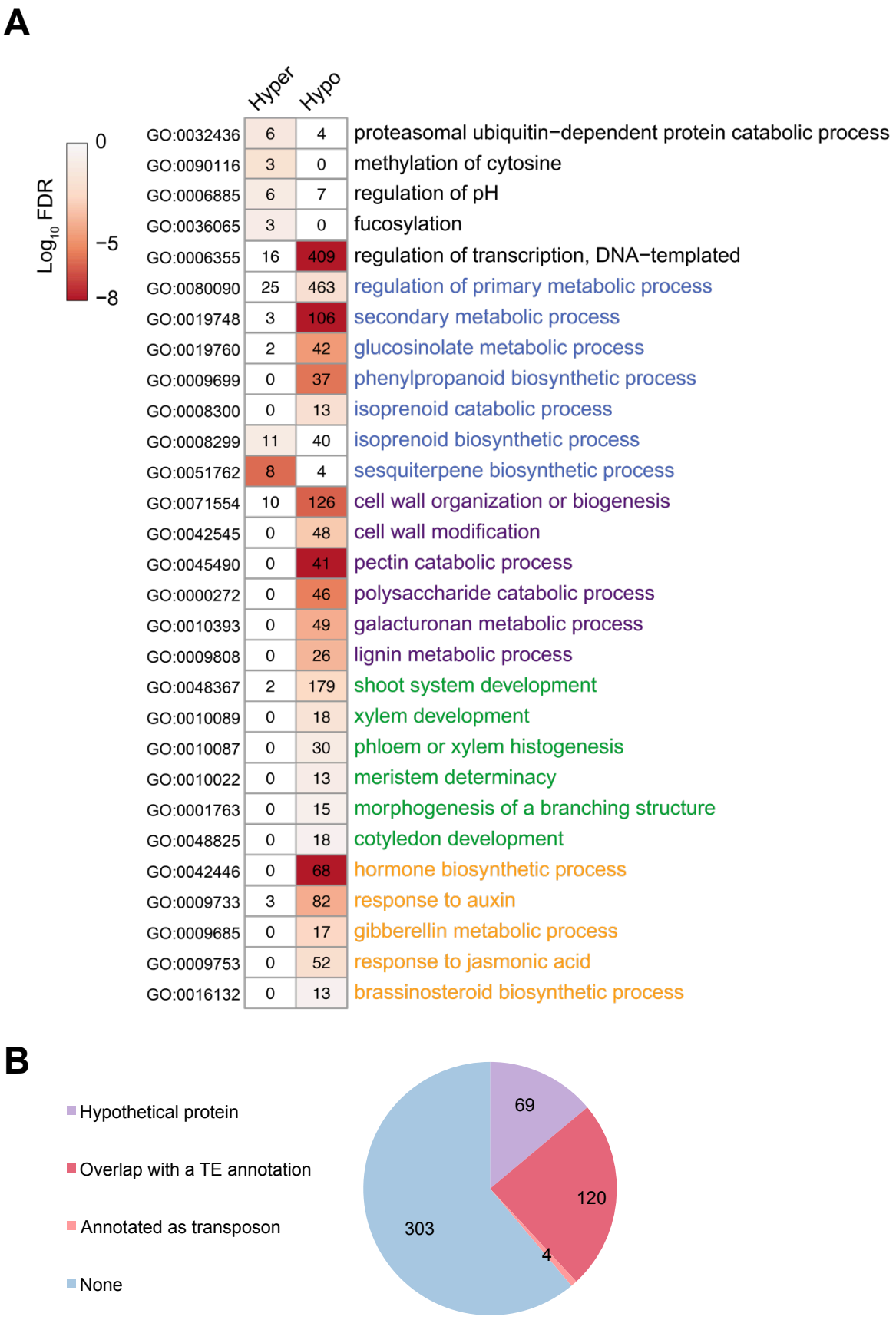


Figure S5. A. Gene ontology analysis of H3K27me3 differentially marked genes in *2h1* plants. A. Association to a significantly over-represented gene function is denoted as a heatmap of false discovery rate (FDR). N=4317 hypo-marked genes; 496 hyper-marked genes. **B.** Number of genes among the 496 hyper-marked genes that either overlap an annotated TE, are annotated as transposons, or are annotated as a hypothetical proteins.

Figure S6

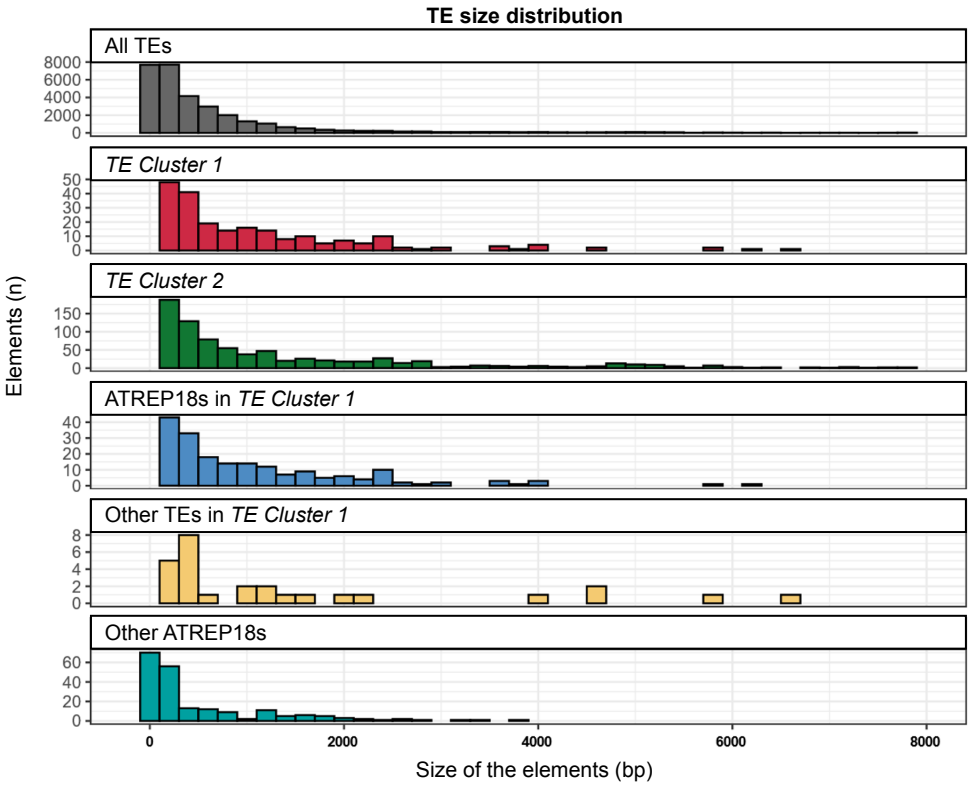


Figure S6. Many *TE cluster 1-2* elements tend to be longer than average TEs.
Size distribution among the repertoire of TEs and repeats analyzed in Figure 3.

Figure S7

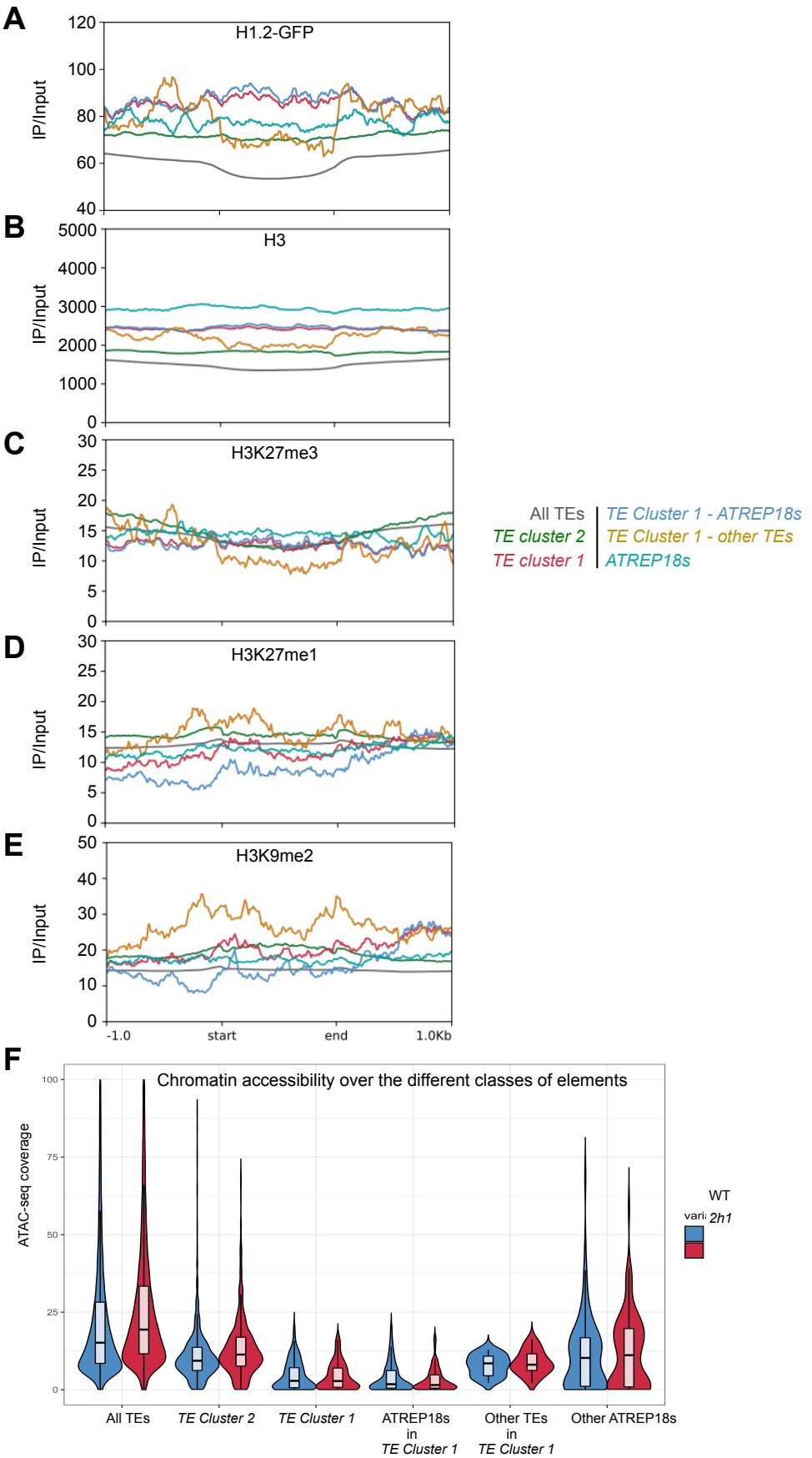


Figure S7. Profiling of representative histone marks on *TE clusters 1-2* and on *ATREP18* elements.
A. H1.2-GFP mean enrichment over the indicated genomic elements shows that *TE cluster 1*, and more generally *ATREP18* elements, display elevated H1 occupancy as compared to other TEs. **B.** H3 profiling indicates that *ATREP18s* and *TE cluster1-ATREP18* elements display elevated nucleosome occupancy. This is consistent with the weak accessibility of these elements determined by ATAC-seq analyses. **C.** Comparison of H3K27me3 profiles indicates that all TE types investigated in this study, including *ATREP18* elements, display similarly low H3K27me3 levels in WT plants. **D.** Profiling of the heterochromatic H3K27me1 mark does not allow identifying a specific predisposition of *TE clusters 1-2* or of *ATREP18* elements for H3K27me1 being used as a substrate for H3K27me3 enrichment. **E.** Profiling of the H3K9me2 mark indicates that *TE cluster 1-2* both display heterochromatic features, which corroborates their elevated nucleosome occupancy and their preferential distribution within pericentromeric regions. All profiles have been generated in this study except H3K27me1 and H3K9me2 (Additional file 4). **F.** ATAC-seq mean read coverage along the repertoire of TEs and repeats indicated.

Figure S8

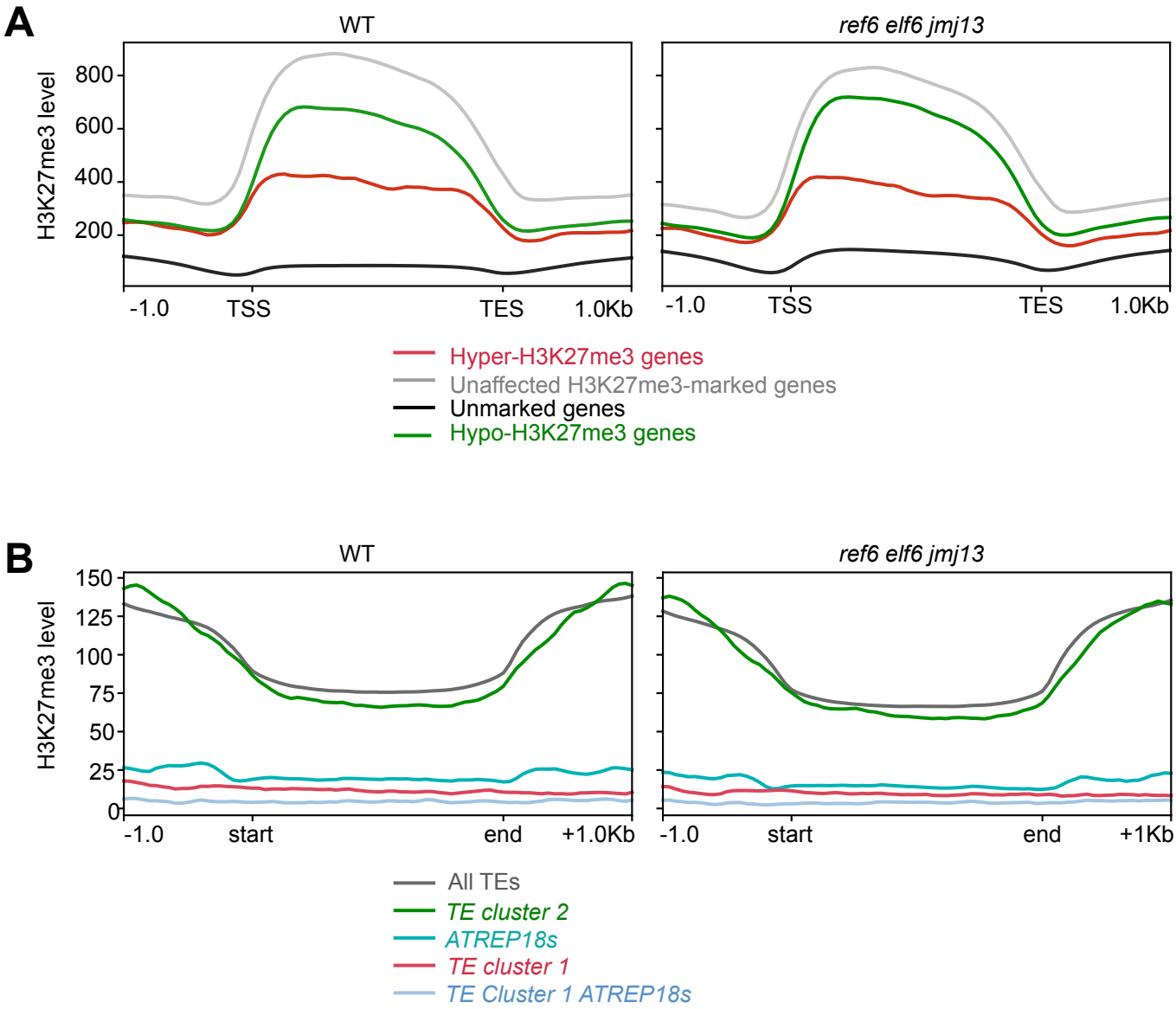


Figure S8. H3K27me3 profiles of the indicated gene (A) and TE (B) sets in WT and *ref6 elf6 jmj13* triple mutant plants impaired in H3K27me3 demethylation. H3K27me3 data were obtained from (Yan et al., 2018).

Figure S9

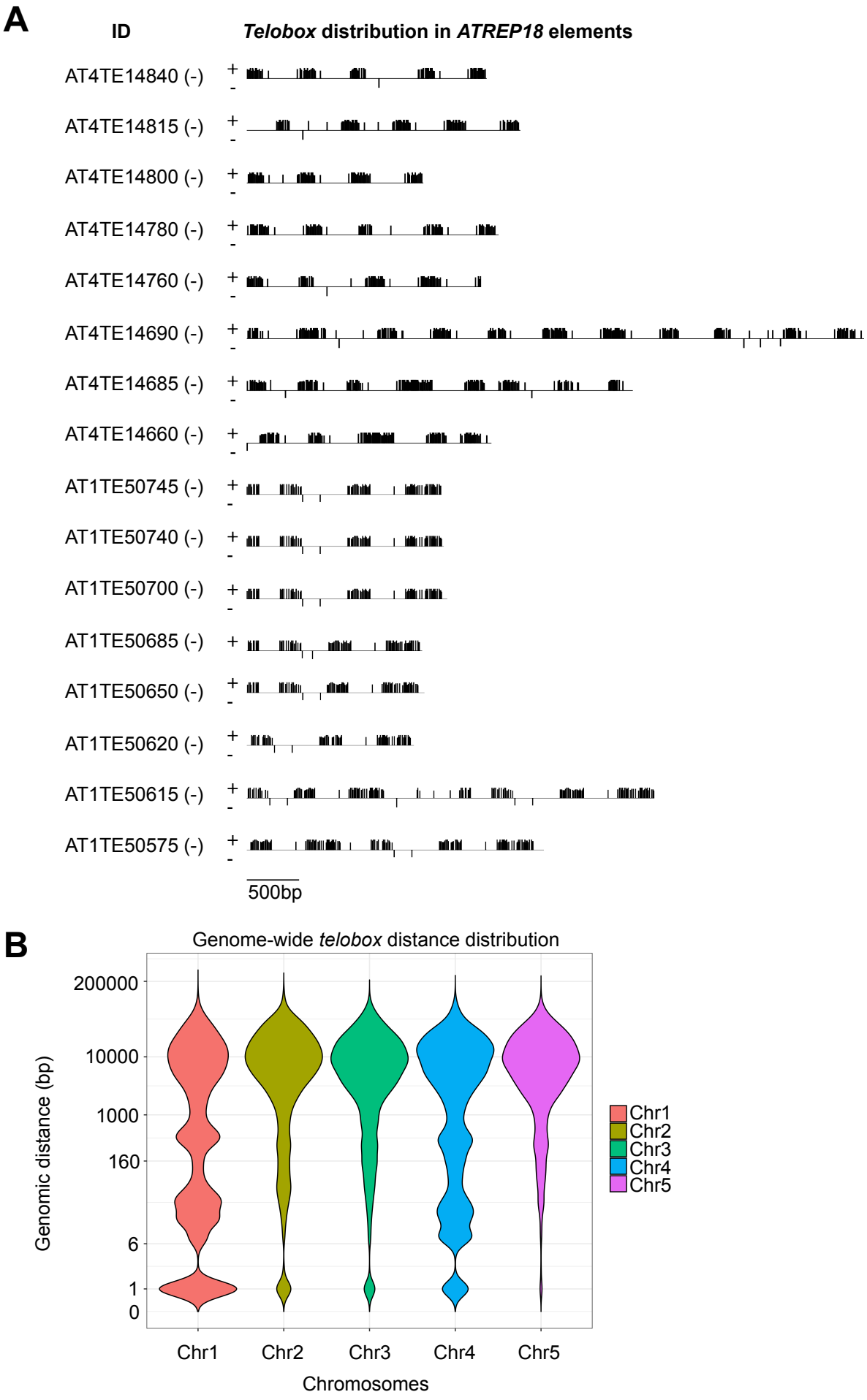


Figure S9. *Telobox* motifs are organized as small clusters in *ATREP18* elements of chromosomes 1 and 4 interstitial telomeric regions. A. *ATREP18* elements in *TE cluster 1* display a particular pattern, with multiple clusters of *telobox* motifs. **B.** Genomic distances between all perfect *telobox* motifs on each chromosome.

Figure S10

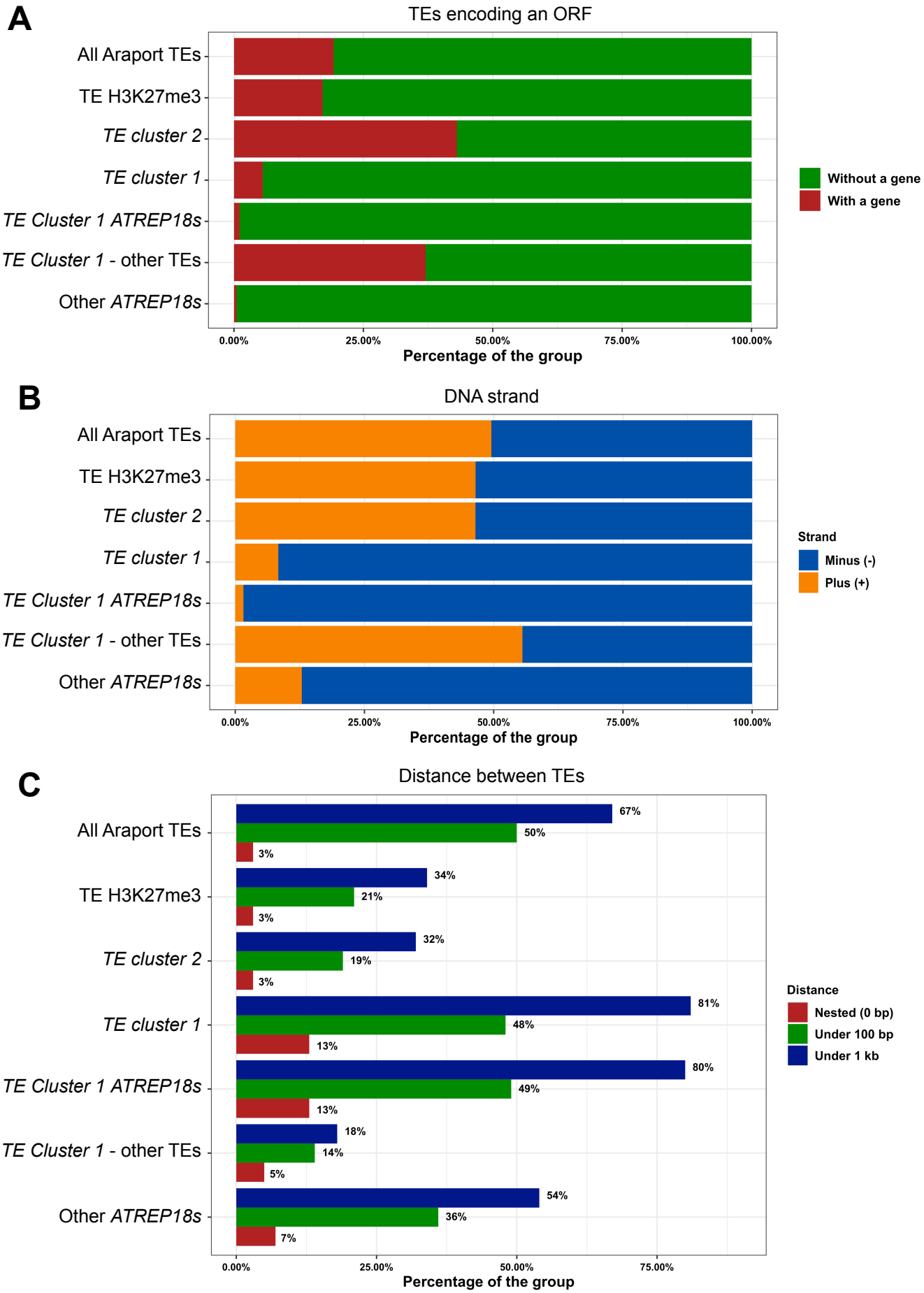


Figure S10. Sequence features of *TE cluster 1-2* elements.
A. Frequency of open reading regions (ORFs) within the indicated sets of TEs. *TE cluster 2*, which is made of a large variety of TE super-families dispersed in the genome, more frequently encodes ORFs compared to the ensemble of all TEs. In contrast, *TE cluster 1*, which contains many pericentromeric *ATREP18* elements, rarely encodes ORFs. **B.** Strand distribution. *TE Cluster 1*, consisting mainly of *ATREP18* elements, is largely strand-specific. **C.** Distribution of different groups of TEs in three classes of distances: nested (0 base pairs), clustered (under 100-bp) or closely located (under 1-kb). Compared to the ensemble of all TEs, *TE cluster 2* elements tend to be dispersed across the genome while, conversely, *TE cluster 1* elements tend to be located in close proximity. This peculiar distribution is largely due to the overwhelming presence of *ATREP18* elements in *TE cluster 1* elements in this group (189 out of 216, 87%).

Figure S11

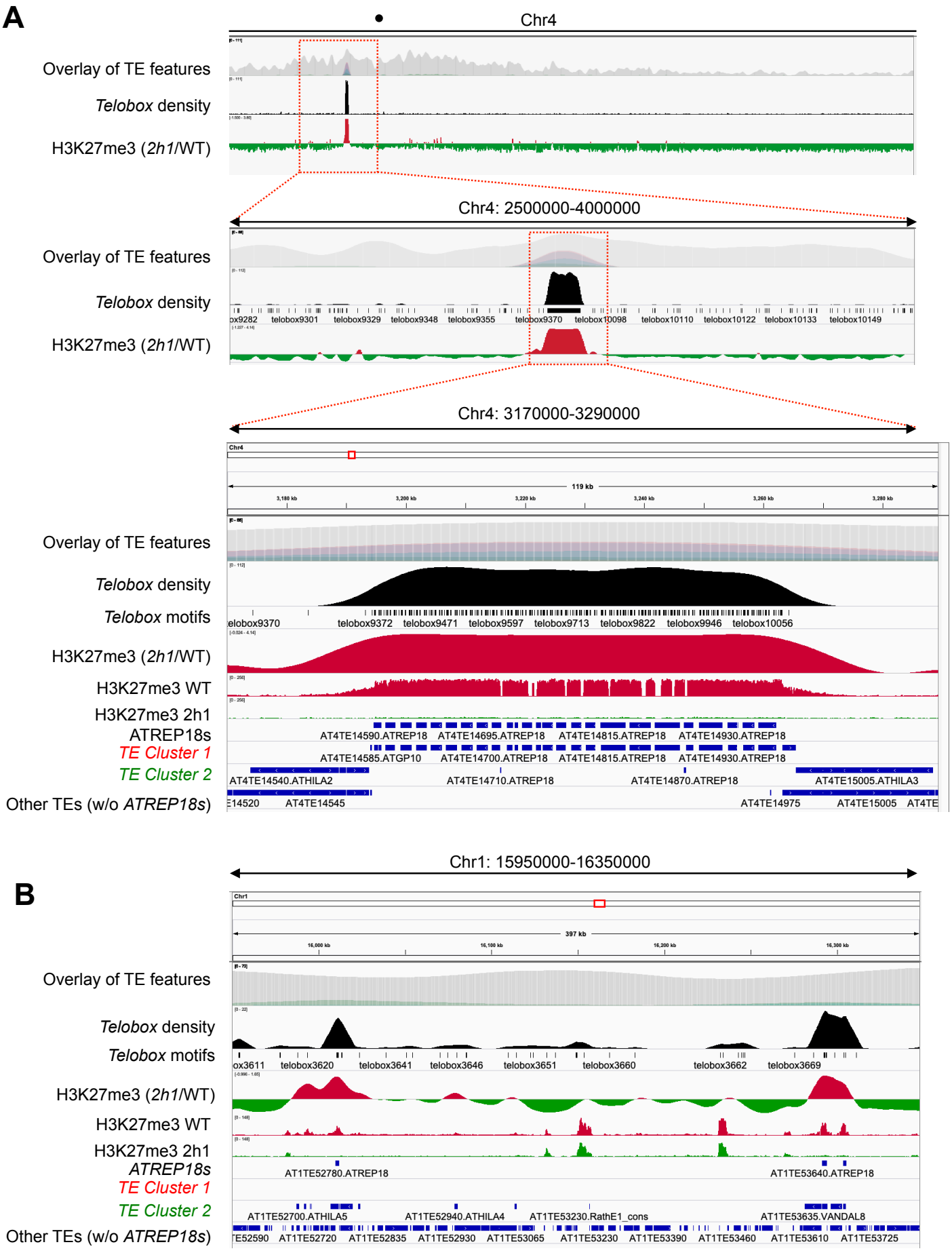
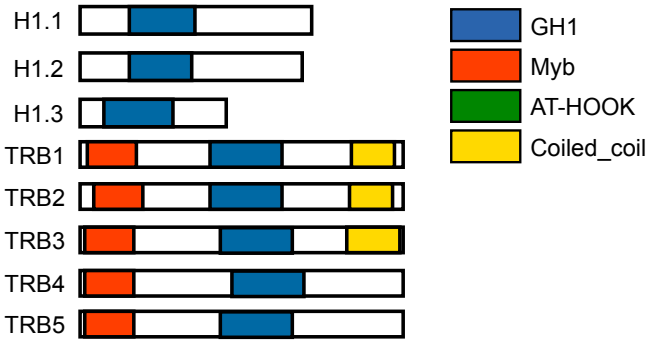


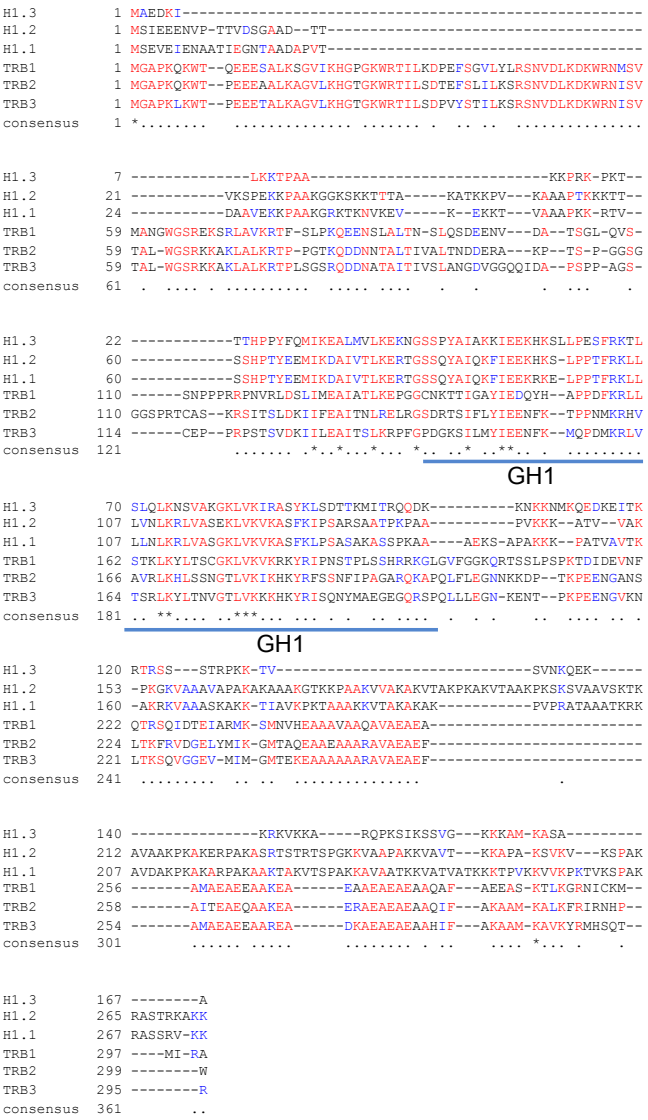
Figure S11. Browser visualization of H3K27me3-enriched repeats in *TE clusters 1-2* and in the *ATREP18* family. **A.** Chromosome 4 distribution of the indicated TEs and *ATREP18* elements together with *telobox* motif distribution. **B.** Close-up view on some interspersed H3K27me3-enriched repeats on chromosome 1 being part of *TE cluster 2*, exemplifying a physical correlation between H3K27me3 enrichment and *telobox*-rich domains in *2h1* pericentromeres.

Figure S12

A



B



C

Identity (%)	H1.1	H1.2	H1.3	TRB1	TRB2	TRB3
H1.1		56.55	23.02	16.57	15.47	16.28
H1.2			22.22	12.78	13.73	13.93
H1.3				11.39	12.26	12.14
TRB1					47.73	47.87
TRB2						65.79
TRB3						

Figure S12. TRB family members display an amino-terminal Myb domain and a central GH1 domain that may explain their association to *telobox*-containing linker DNA. A. Schematic representation of H1 and TRB proteins with all identified protein domains. **B.** ClustalW protein sequence alignment of TRB1, TRB2 and TRB3 proteins with the three H1 variants showing the relative conservation of a central globular H1 domain. **C.** Amino-acids sequence identify between TRB1, TRB2 and TRB3 proteins and the three H1 variants.

Figure S13

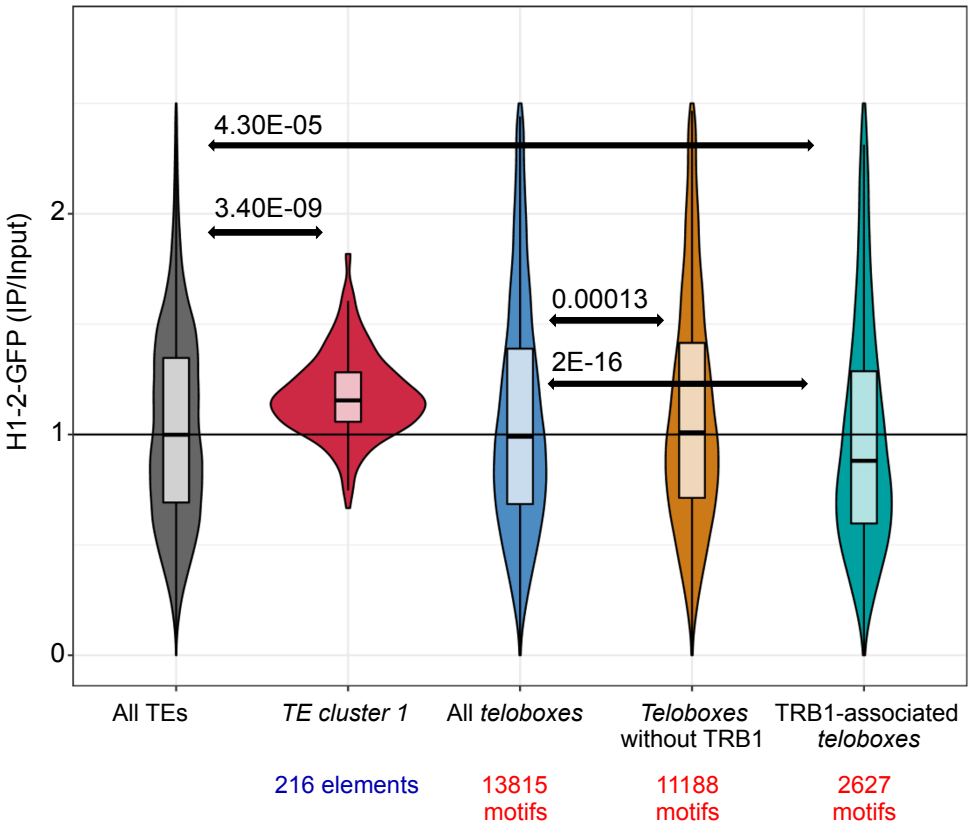


Figure S13. *TE cluster 1* displays elevated H1 level as compared to the ensemble of *Arabidopsis* TEs and to other *telobox*-containing regions. In agreement with the proposed mutually exclusive binding of H1 and TRB1 over *teloboxes*, the set of *teloboxes* with a known TRB1 peak from ChIP-seq data (Schrumpfová et al., 2014) displays significantly less H1 occupancy than other *teloboxes*. Statistical significances of differences between the mean levels were assessed using a Wilcoxon signed-rank test.

Figure S14

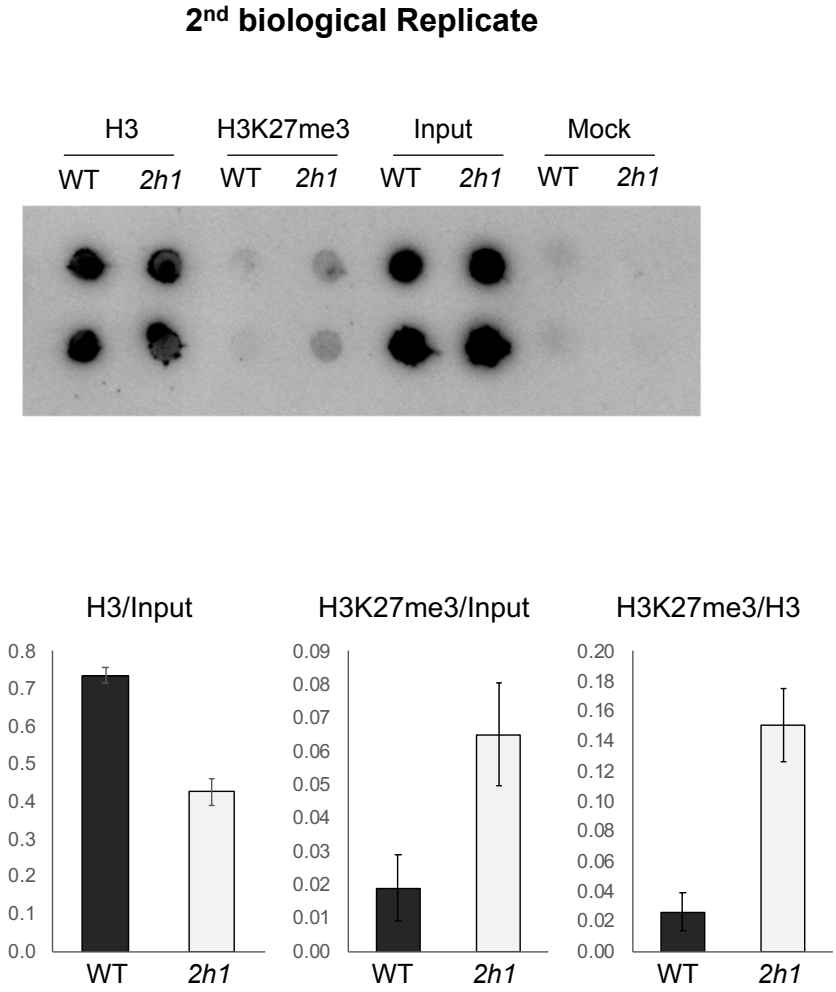


Figure S14. Independent biological replicate of H3K27me3 and H3 ChIP-hybridization to telomeric probes completing Figure 6A.

Figure S15

A

	Genotype	Sample type	Enzyme	Total sequenced reads	Dangling end	Religation pairs	Self circle	% PCR duplicates	Valid interactions
This study	WT Col-0	Cotyledons	DpnII	403819367	74264813	17814786	412682	26	250301412
	2h1	Cotyledons	DpnII	350357481	58730097	15419444	324512	24	206805943
Liu et al., 2016	WT Col-0	Entire seedlings	DpnII	143406012	35827356	8982912	1110590	17	69657043
Liu et al., 2016	WT Col-0	Entire seedlings	DpnII	160112293	31038473	8095915	873572	19	80536362
Wang et al., 2016	WT Col-0	Entire seedlings	DpnII	247618590	32238398	9134591	1914560	29	146107616
Wang et al., 2016	WT Col-0	Entire seedlings	DpnII	239477694	42804562	12643414	726839	25	141113166

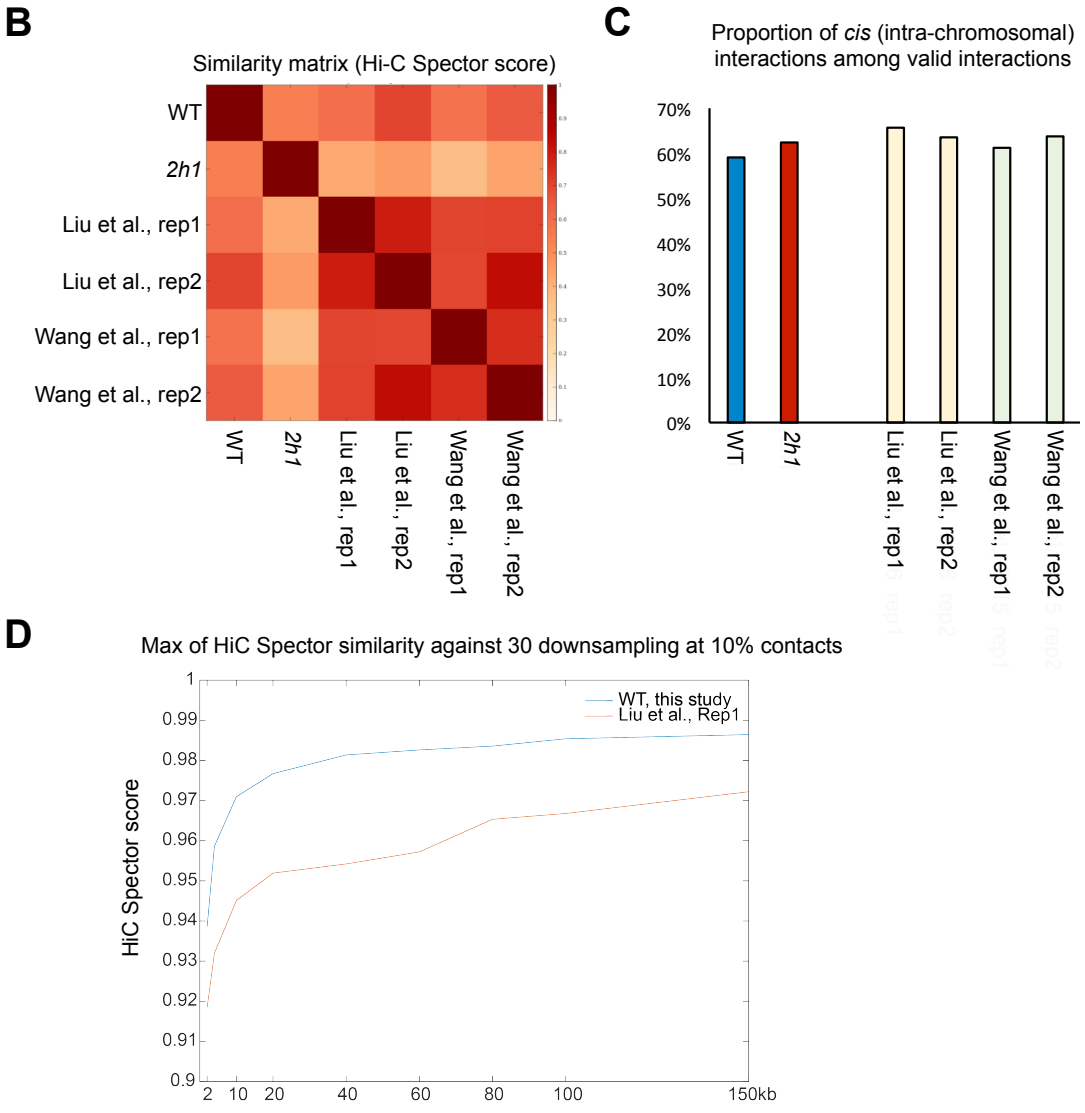


Figure S15: Hi-C mapping and similarity index quality.
A. Mapping results of Hi-C libraries and comparison with re-processed Hi-C datasets from Liu et al., (2016) and Wang et al., (2015). **B**. Heatmap analysis of similarity among WT and *2h1* mutant datasets with publicly available *Arabidopsis* Hi-C datasets generated with the DpnII enzyme. Similarity score was calculated using Hi-C Spector score and 100 kb bins. **C**. Comparison of intra-chromosomal reads over the total number of valid interactions among our samples and publicly available data. This index is positively correlated with library quality for *in situ* Hi-C (Sun et al., 2020). Similar values were scored for processed samples (WT and *2h1)* and published DpnII-based Hi-C datasets. **D.** Estimation of the resolution achieved in this Hi-C study. The curves show the similarities between this study and another high-quality DpnII-based dataset by Liu et al. (2016) using HiC Spector score of chromosome 1 contact map against himself down-sampled at 10% of contact as in (Carron et al., 2019). We computed 30 down-samples at a resolution of 2kb and computed HiC Spector similarity against the original Hi-C map with a resolution from 2kb to 150kb. For each resolution, we only took the maximum HiC Spector score of the 30 down-samplings. Blue curve, analysis of our WT Hi-C data. Red curve, analysis of replicate 1 from Liu et al. (2016).

Figure S16

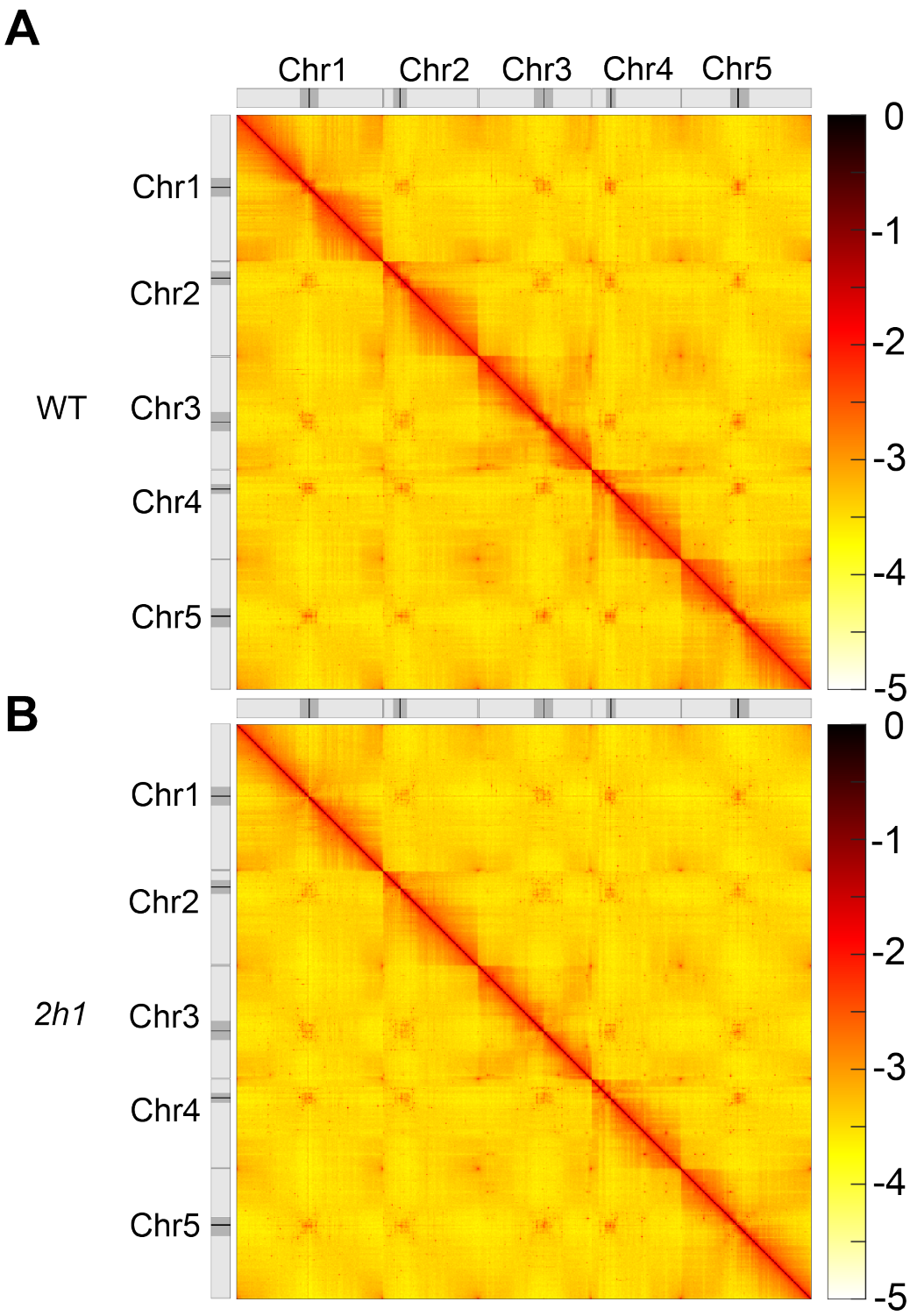


Figure S16. Heatmap of Log_{10} contact count of WT (A) and *2h1* (B) samples, normalized by SCN on the whole genome at a 10 kb resolution.

Figure S17

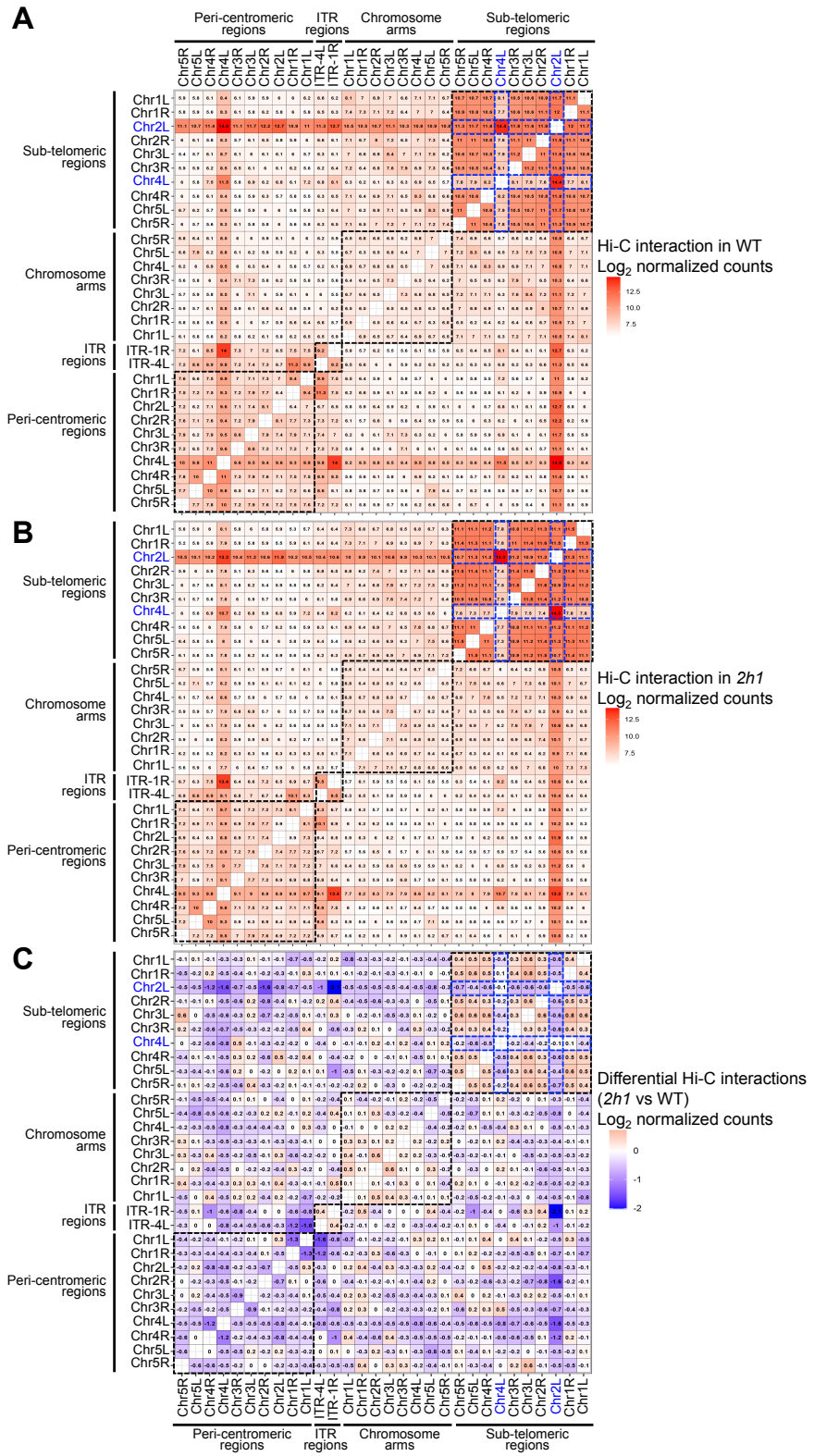


Figure S17. Heatmap showing the frequency of interaction between distinct regions of the genome in WT (A) and *2h1* mutant plants and the difference between them. We probed four groups of regions: 1) ITR-1R and ITR-4L coordinates, s2) sub-telomeric regions defined as the 100 kb terminal chromosomal regions that neighbor either the telomeres or NOR2 and NOR4, 2) pericentromeric regions represented by 100-kb segments located at 1 Mb from centromeres, and 4) distal arm 100-kb regions located at 5 Mb from the telomere positions. The sub-telomeric regions of chromosome 2 and 4 left arms (Chr2L and Chr4L) are therefore linked to NOR2 and NOR4, respectively (blue label). **A-B.** As indicated by dashed lines, regions belonging to the same group tend to interact frequently among themselves in both WT and *2h1* mutant plants. ITRs tend to interact between themselves or more generally with their surrounding regions as ITR-1R frequently interacts with chromosome 4L pericentromere, and conversely ITR-4L frequently interacts with chromosome 1R pericentromere. In contrast, ITRs poorly establish interactions with all the other probed groups. Besides establishing exceptionally frequent mutual interactions between them, NOR2 and NOR4-asociated Chr2L and Chr4L regions also show frequent interactions with ITR-4L and with the neighboring pericentromeric regions of Chr4L. **C.** Comparison of WT and *2h1* plants shows that topological interaction patterns are specifically altered in *2h1* mutant plants. Except for the NOR-associated Chr2L and Chr4L, in the absence of H1, inter-telomere interactions tend to increase among them and with the H3K27me3-enriched ITRs. Similarly, interactions among the ITRs are frequent in the mutant line. In contrast, NOR-associated regions are subjected to less frequent interactions with the other sub-telomeric regions, with the pericentromeres of chromosome 2 and 4, and to a lower extent with all other regions analyzed in *2h1* plants. Similarly, interactions between ITRs and the pericentromeres are greatly reduced in the mutant. Distal arm regions used as control in these analyses expectedly display no clear trend with a mix of positive and negative variations. Interaction frequencies are expressed as logarithm of the observed read pairs (A and B) or logarithm of their ratio (C) normalized for region size. NOR positions are indicated in blue.