**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]Cytometry Facility, Imagerie-Gif, Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, 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

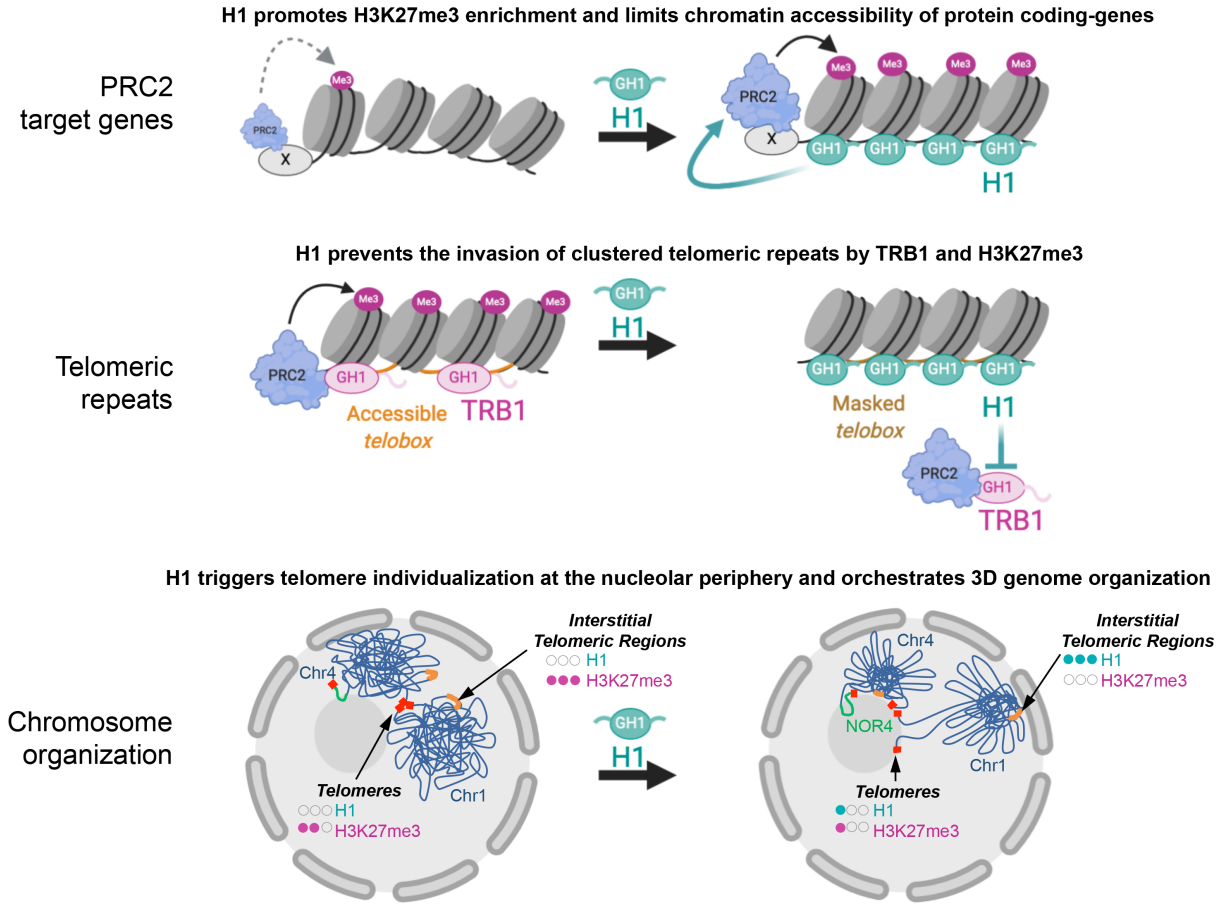
* Correspondence:
clara.bourbousse@bio.ens.psl.eu
barneche@bio.ens.psl.eu

**Running title:** H1 and H3K27me3 interplay in *Arabidopsis*

**Keywords:** *Arabidopsis*, Linker histone H1, H3K27me3, *Polycomb*, Chromatin accessibility, 3D chromosome organization, Telomere, Telomere Repeat Binding (TRB) protein

**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, interplays of H1 with chromatin factors along the epigenome landscape are just starting to emerge. Here, we report that *Arabidopsis* H1 occupies and favors both chromatin compaction and H3K27me3 enrichment on a majority of *Polycomb*-target protein-coding genes. In contrast, H1 prevents H3K27me3 accumulation on telomeres and pericentromeric interstitial telomeric repeats (ITRs) while orchestrating long-distance interactions regulating the 3D organization of these chromosome regions. Mechanistically, H1 prevents ITRs from being invaded by Telomere Repeat Binding 1 (TRB1), a GH1-containing telomere component with extra-telomeric functions in *Polycomb* recruitment. Based on these findings, we propose that H1 represses H3K27me3 accumulation on large blocks of telomeric repeats by antagonizing TRB1 association to linker DNA, conferring to linker histones an additional and sequence-specific role in modulating H3K27me3 epigenome homeostasis.



## 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. H1 also 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 (reviewed in Bonev and Cavalli, 2016), in which both H1 and *Polycomb* complexes play essential roles (Feng et al., 2014; Geeven et al., 2015; Grob et al., 2014; Liu et al., 2016; Sexton et al., 2012). 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, mammalian 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 itself (Margueron et al., 2008). Specific functional interplays between PcG and H1 in chromatin compaction have also emerged (Yuan et al., 2012). In human, the preferential interaction of H1.2 with H3K27me3 nucleosomes promotes chromatin compaction (Kim et al., 2015), while, conversely, human and mouse PRC2 complexes display substrate preferences for H1-enriched chromatin fragments, their *in vitro* activity being more stimulated on dinucleosomes than on mononucleosomes (Martin et al., 2006; Willcockson et al., 2020). In agreement with these findings, mouse H1 has just been identified as a critical regulator of H3K27me3 enrichment over hundreds of PRC2 target genes (Willcockson et al., 2020; Yusufova et al., 2020). Interestingly, chromosome conformation capture (Hi-C) analysis of haematopoietic cells (Willcockson et al., 2020), germinal centre B cells (Yusufova et al., 2020), and embryonic stem cells (Geeven et al., 2015) showed that H1-mediated chromatin folding in the nuclear space triggers distinct genome topologies in diverse cell types or during differentiation in mammals. Still, a potential sequence specificity of these events and their conservation through evolution have not yet been addressed.

In *Arabidopsis thaliana*, the H1.1 and H1.2 canonical linker histone variants display similar chromatin association properties and represent the full H1 complement in most somatic cells while a third

and atypical H1.3 variant is only expressed under stress conditions or in a few cell types (reviewed in Kotliński et al., 2017; Over and Michaels, 2014; Probst et al., 2020). H1.1 and H1.2, hereafter referred to as H1, are enriched over heterochromatic transposable elements (TEs), which also display high levels of cytosine methylation, of nucleosome occupancy, and of heterochromatic histone modifications such as H3K9me2 (Choi et al., 2019; Rutowicz et al., 2015). 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 mammals, incorporation of *Arabidopsis* H1 is thought to dampen transcription elongation, an effect that, in plants, 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, the H1 complement is massively degraded in the vegetative pollen cell nucleus (He et al., 2019; Hsieh et al., 2016) and during the formation of *Arabidopsis* spore mother cells (SMCs), a depletion that coincides with heterochromatin loosening and a reduction in most H3K27me3 signals in SMC nuclei (She et al., 2013; She and Baroux, 2015). Hence, despite evidence that control of H1 abundance is used by plants for a global reprogramming of their epigenome, there is currently no information on the consequences of H1 depletion on the H3K27me3 chromatin landscape and, more generally, in the dynamic equilibria modulating chromatin compaction and accessibility in plants.

Using cytogenetic and bulk analyses of chromatin, we previously reported that somatic cells of *Arabidopsis H1* knockout plants not only display dramatic defects in heterochromatin compaction but also present globally reduced H3K27me3 abundance whilst, oppositely, a few discrete subnuclear foci of undetermined nature display increased H3K27me3 signals (Rutowicz et al., 2019). Here, we report that H1 is required for H3K27me3 enrichment and low accessibility on a majority of the PRC2 target genes while preventing the accumulation of this mark and influencing the 3D organization of telomeres and pericentromeric interstitial telomeric regions. We explored the specificity of these chromatin alterations through the involvement of Telomere Repeat Binding 1 (TRB1), a GH1-containing histone-related protein that displays extra-telomeric roles in PRC2 recruitment on telomeric motifs (Schrumpfová et al., 2016; Xiao et al., 2017; Zhou et al., 2018). Collectively, our findings led us to establish that H1-mediated chromatin compaction orchestrates *Arabidopsis* chromosomal organization and contributes to the control of H3K27me3 homeostasis between structurally distinct regions in a sequence-specific manner, notably by antagonizing DNA binding of TRB proteins.

## Results

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

To assess the relationships between H3K27me3, chromatin accessibility and the H1 landscapes, we first compared the genomic distribution of H3K27me3 with that of H1.2, the most abundant canonical H1 variant. To maximize H1 ChIP-seq specificity, we used an H1.2 GFP-tagged version 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 quasi-exclusive carriers of H3K27me3 in *Arabidopsis* (Roudier et al., 2011; Sequeira-Mendes et al., 2014; Wang et al., 2015), showed that H1 is highly abundant on H3K27me3-marked gene bodies, especially towards their 5' region in comparison to non-marked genes (Figure 1A), and this independently from variations in histone H3 profiles used as a proxy of nucleosome occupancy (Figure S1A).

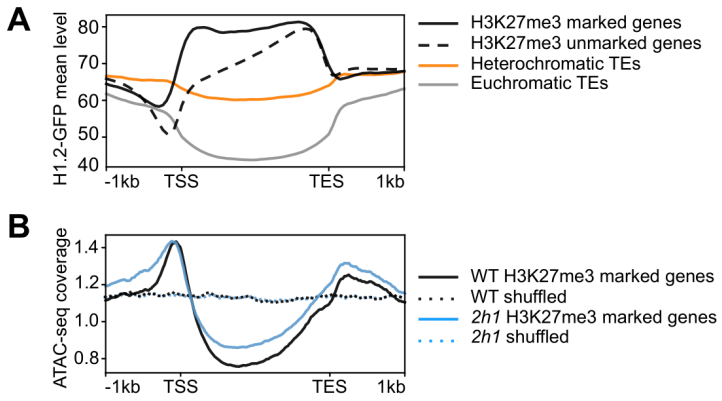



**Figure 1. H1.2-GFP is enriched over PRC2-target chromatin and contributes to restrain its accessibility. A.** H1.2-GFP mean level over different genetic contexts. H3K27me3-marked genes defined in our ChIP-seq analysis (n=7542) are compared to all other annotated protein-coding genes. Heterochromatic 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.

Examination of transcription start sites (TSS) confirmed that H1 is globally more abundant on H3K27me3-marked genes than on genes with hallmarks of active Pol II initiation (H3K4me3), elongation (H2Bub), or having high transcript levels (Figure S1A). H1 also was globally much more present on H3K27me3-marked genes than on heterochromatic H3K9me2-rich transposable elements (TEs) (Figure 1A), which are themselves H1-rich and heavily condensed (Choi et al., 2019). This genome-wide enrichment suggests local interplays between H1 and H3K27me3 on protein-coding genes.

We examined chromatin accessibility of H3K27me3-marked genes by Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) analysis of 4C nuclei, corresponding to the most abundant ploidy level in our samples. In WT plants, H3K27me3-marked genes displayed low chromatin accessibility as compared to expressed genes (Figure S1B), which typically show a sharp ATAC peak at their TSS in such analyses (Lu et al., 2016). We similarly examined *h1.1h1.2* (*2h1*) double mutant plants to test whether H1 contributes to this low accessibility. In contrast to transcriptionally active genes, H3K27me3-marked genes display no change in the TSS accessibility peak upon H1 depletion but they are globally more accessible along their body as well as their promoter and terminator domains (Figure 1B and S1B). Hence, H1 strongly occupies H3K27me3-marked genes where it could favor low chromatin accessibility.

### H1 is necessary for H3K27me3 enrichment at a majority of PRC2 target genes

As a first assessment of functional links between H1 and H3K27me3, we profiled the H3K27me3 landscape upon H1 depletion. Considering that H3K27me3 bulk levels are ~2-fold lower in immunoblot and immunocytological analyses of *2h1* seedlings nuclei (Rutowicz et al., 2019), we employed a spike-in ChIP-seq with reference exogenous genome (ChIP-Rx) approach using a fixed amount of Drosophila chromatin added to each sample prior to immunoprecipitation and subsequently quantified in sequenced

input and immunoprecipitated samples (as in Nassrallah et al., 2018) (Figure S2 and Additional File 1). Among the 7542 genes significantly marked by H3K27me3 in WT plants, 4317 (~55%) of them displayed lower levels of the mark in *2h1* plants while only 496 had the opposite tendency (Figure 2A-C, S3 and Additional File 1). Hence, the decreased levels of H3K27me3 in *2h1* seedlings previously identified by immunoblotting and immunocytology (Rutowicz et al., 2019) result from a widespread effect of H1 on PRC2 target genes.

We first envisaged that local H1 stoichiometry may contribute to determine PRC2 marking in *cis* as recently reported in mouse (Willcockson et al., 2020). We cannot rule out that a similar mechanism is at play in *Arabidopsis* but, oppositely, we found that H3K27me3 hypo-marked genes tend to display lower H1 levels in WT plants than unaffected ones (Figure 2D). Hence, in the absence of direct correlation underlying a selective influence of H1 on genes, we envisaged a sequence-dependent mechanism and explored it by examining sequence motifs in the promoters of H3K27me3 hypo/hyper-marked gene sets. This did not reveal any over-represented sequences in the promoters of the thousands hypo-marked genes. In contrast, we found three prevalent motifs in the 496 promoter sequences of the hyper-marked gene set (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 (*PREs*) in *Arabidopsis* (Xiao et al., 2017; Zhou et al., 2018). In total, 87 of the 496 hyper-marked genes contain one or more *telobox* motifs in their promoters (Additional File 1).

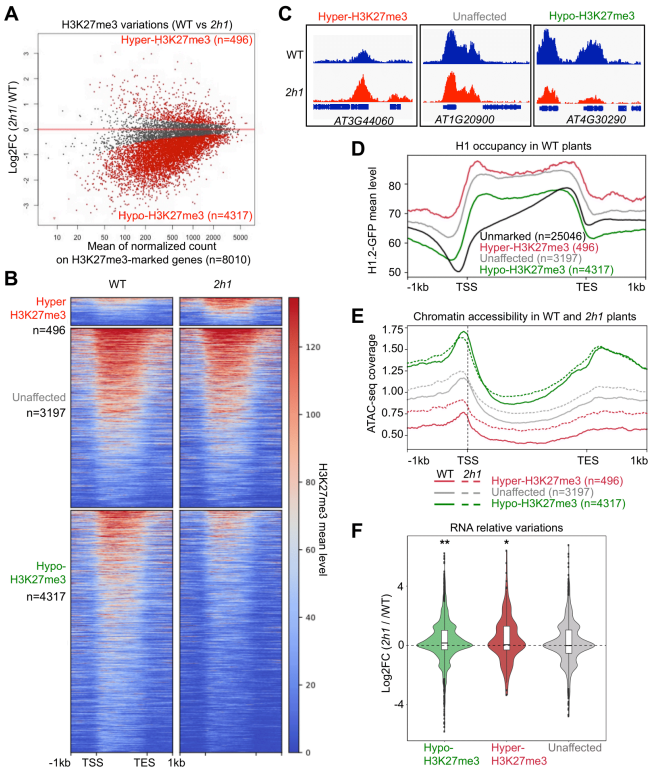



**Figure 2. H1 influences H3K27me3 marking, chromatin accessibility and expression of PRC2-target genes. A.** Identification of differentially marked genes using spike-in normalized DESeq2 analysis shows that *H1* loss-of-function triggers low H3K27me3 levels over a majority of the PRC2 target genes. 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 differences in H3K27me3 levels after spike-in normalization identified in (A) and ranked in each group according to mean H3K27me3 levels. **C.** H3K27me3 profile on representative genes of the same three gene sets identified in (A) exemplifying the general tendency of PRC2-target genes to keep a weak H3K27me3 domain in *2h1* plants. **D.** H1.2-GFP profile on the same three H3K27me3-marked gene sets. **E.** ATAC-seq analysis of the three H3K27me3-marked gene sets. 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 $Log_2$ 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 over a large set of genes, we examined in our ATAC-seq data whether they had distinct chromatin accessibilities as compared to unaffected genes. In line with their moderate H1 occupancy, H3K27me3 hypo-marked genes were globally more accessible in WT plants than unaffected or hyper-marked genes (Figure 2D and E). We

then tested chromatin accessibility of these gene categories in *2h1* nuclei. Hypo-marked gene bodies were slightly more accessible in the mutant line (Figure 2E), thereby correlating with local reduction in H3K27me3. Interestingly, accessibility of hyper-marked genes was also globally increased in *2h1* plants despite a clear gain in H3K27me3. These discrepancies suggested that mostly H1 incorporation and not variations in H3K27me3 influence the accessibility of PRC2-target genes.

**H1 contributes to define expression of PRC2 target genes**

To get more insights into the biological role of H1 on H3K27me3-marked genes, hereafter referred to as PRC2-target genes, we compared their transcript levels in WT and *2h1* plants. Similarly to previous reports (Choi et al., 2019; Rutowicz et al., 2019), our RNA-seq analysis confirmed that *H1* loss-of-function triggers only minor gene expression changes in *Arabidopsis* seedlings (Additional File 2), but focusing on the set of H3K27me3 hypo-marked genes showed a significant upregulation of this gene repertoire in *2h1* plants (Figure 2F). Functional categorization of the most significantly misregulated ones identified an over-representation of genes involved in transcriptional regulation, meristem maintenance, cell wall organization and vascular development (Figure S5A). These categories are consistent with the repression of these biological processes by PRC2 (de Lucas et al., 2016) and with the subtle PRC2-like phenotypes of *H1* mutant plants (Rutowicz et al., 2019). Reciprocally, a survey of the hyper-marked genes identified a large proportion of TE-like features, as 25% (124/496) of them overlap with a TE annotation or are annotated as transposon genes (Figure S5B, Additional File 1). Collectively, we identified a positive effect of H1 on H3K27me3 enrichment and chromatin compaction; effects that may jointly influence expression. This trend is contrasted by H1 preventing H3K27me3 accumulation at a minority of H1-rich and poorly expressed genes that frequently display TE features.

**H1 prevents H3K27me3 invasion over a specific family of heterochromatic repeats**

Considering the unexpected *de novo* marking of H3K27me3 on several TE-related genes in *2h1* plants, we extended our analysis to all TEs that, in *Arabidopsis*, usually lack H3K27me3 at baseline. 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 (Figure 3A). We clustered them into two groups, *TE cluster 1* (n=216) with high enrichment of H3K27me3, and *TE cluster 2* (n=850) with modest 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 corresponding to *ATREP18* elements (Figure 3B). This is a strong over-representation as, in total, *TE cluster 1* and *2* comprise 60% of the 391 *ATREP18s Arabidopsis* elements (189 and 47, respectively), including many of the longest units (Figure S6). Comparison of H3K27me3 variation with local H1 occupancy showed that *ATREP18* elements stand out from the general population of TEs by their outstanding gain of H3K27me3 in *2h1* plants (red dots in Figure 3C). *TE cluster 1,* and more generally *ATREP18* elements, display heterochromatic properties with H3K9me2 marks (Figure S7A-D), elevated nucleosome and H1 occupancies (Figure 3D) as well as an extremely low chromatin accessibility as compared to the majority of TEs (Figure 3E and S7E). Taken together, these observations indicate that H1 has a negative effect on H3K27me3 accumulation over a set of H1-rich, heterochromatic, and highly compacted repeats, which contrasts with H1's positive influence on H3K27me3 marking over thousands of PRC2-target genes.

While TEs are, on average, more accessible in *H1* loss-of-function mutant plants (Choi et al., 2019; Rutowicz et al., 2019), ATAC-seq analysis of *TE cluster 1* and all *ATREP18* repeats showed that chromatin of these elements remained very poorly accessible in *2h1* plants (Figure 3E and S6D). This observation indicates that chromatin "inaccessibility" of *ATREP18* elements is H1-independent or that H1 depletion is compensated by other mechanisms, such as the strong H3K27me3 local enrichment.

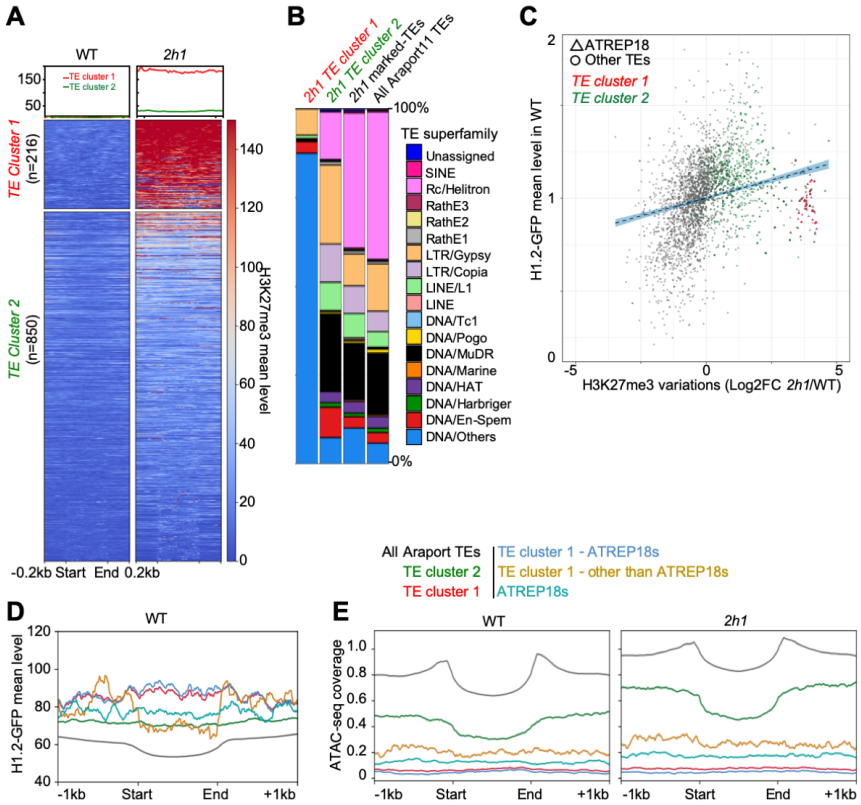



**Figure 3. H1 is required to avoid elevated H3K27me3 levels on a specific set of repeated elements. A.** H3K27me3 profiles over all *TE Cluster 1* and *2* elements. Hyper-marked TEs were clustered into two groups according to H3K27me3 levels after spike-in normalization; defining *TE Cluster 1* and *2*, 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 mean 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

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

To further explore how H1 selectively contributes to prevent H3K27me3 marking on *TE cluster 1-2*, we first envisaged the H3K27me1 heterochromatic mark as a preferred substrate for H3K27 trimethylation. However analysis of public datasets (Ma et al., 2018) showed that, as compared to other TEs, H3K27me1 is neither particularly abundant on *TE cluster 1-2* nor on *ATREP18* elements (Figure S7C). 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 selective H3K27me3 removal on *TE cluster 1* elements in WT plants (Figure S8), altogether ruling out a specific role for H1 in promoting H3K27me3 demethylation at these loci.

We subsequently searched for over-represented DNA motifs in *TE cluster 1* elements. Out of three 5-9 bp motifs identified, the most significantly enriched sequence corresponds to the *telobox* (ACCCTAA) motif (Figure 4A), as formerly found in TE-related genes gaining H3K27me3 in *2h1* plants (Figure S4). In total, 7328 *teloboxes* were found in 195 of the 216 *TE cluster 1* elements (Additional File 1). *Vice versa*, searching this motif in all 391 annotated *ATREP18* elements identified more than ten thousand motifs frequently organized as small clusters (Figure S9), which correspond to a ~100-fold over-representation as compared to other TEs (Figure 4B – note the two-colored scales). Besides these motifs, *ATREP18s* display neither typical TE functional features nor a predicted protein-coding region (Figure S10A). They are principally oriented on the minus DNA strand (Figure S10B) and most frequently located in close vicinity, nearly 90 % of them being positioned within 1 kb of each other (Figure S10C). Consistent with this spatial proximity, analysis of chromosomal distribution of *ATREP18s* revealed their concentration within two regions of ~355 kb and ~72 kb on chromosomes 1 and 4, respectively, which both colocalize remarkably well with outstanding H3K27me3 gain in *2h1* plants (Figure 4C and S11).

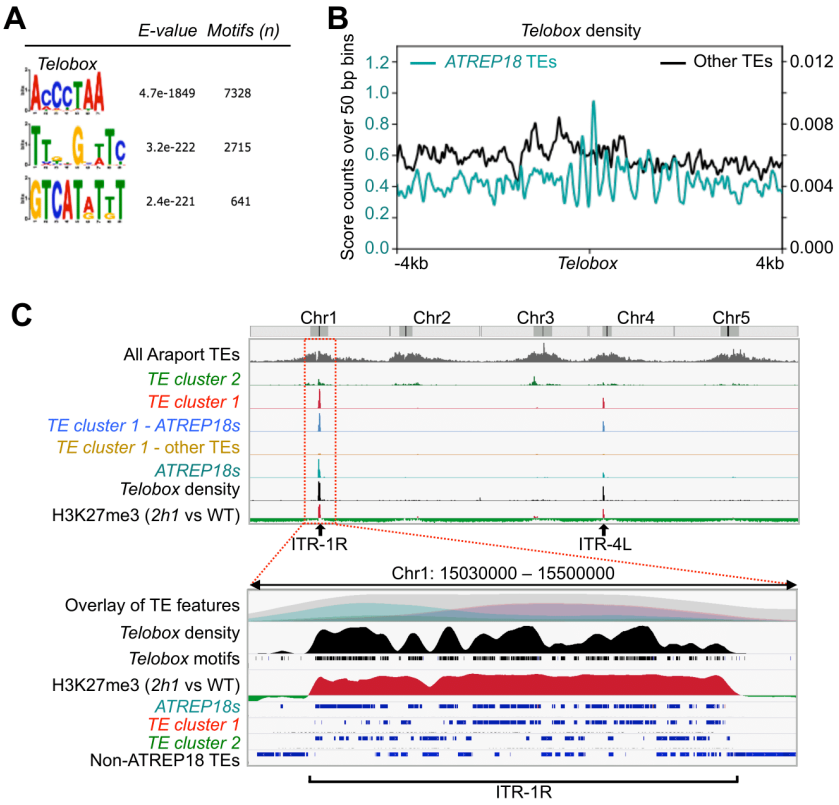



**Figure 4. Two pericentromeric ITR blocks, altogether spanning ~430 kb, are subject to massive H3K27me3 enrichment in *2h1* plants. A.** Prevalent sequence motif over-represented in *TE cluster 1* elements (E-value<1e-220). E-values were calculated against all Araport11 TE sequences. **B.** *ATREP18* 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 from *TE Cluster 2* are shown in Figure S11.

These specific features led us to consider these large blocks as two of the nine *Arabidopsis* genome loci previously identified as Interstitial Telomeric Repeat regions or ITRs (Uchida et al., 2002; Vannier et al., 2009), which comprise a mix of perfect and degenerated telomeric repeats (Schrumpfová et

al., 2019). Altogether, 95 % of *TE Cluster 1* elements are contained within two ITR blocks embedded in pericentromeric regions of chromosomes 1R and 4L, hereafter referred to as ITR-1R and ITR-4L (Figure 4C). Ectopic H3K27me3 deposition was also found on interspersed *TE cluster 2* elements located in pericentromeric regions of the five chromosomes outside 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 Telomere Repeat Binding 1 protein from ITR blocks**

With the aim of assessing the molecular mechanisms that selectively drive PRC2 activity on ITRs in the absence of H1, 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 to maintain telomere length (Schrumpfová et al., 2014). The Myb domain of TRBs 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 onto gene promoters bearing telomeric *telobox* motifs (Schrumpfová et al., 2016; Zhou et al., 2016), notably through the recruitment of PRC2 (Xiao et al., 2017; Zhou et al., 2018). Interestingly, despite their low protein sequence similarity to 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 chromatin incorporation of these two categories of GH1-containing proteins might modulate PRC2 recruitment on ITRs.

To explore this possibility, we first examined the genomic distribution of TRBs with regard to H1 and *telobox* motifs using available TRB1 ChIP-seq data (Schrumpfová et al., 2016). This profiling analysis confirmed that TRB1 peaks were expectedly centered on *telobox* motifs (Schrumpfová et al., 2016) and that H1 is globally enriched over *TE Cluster 1* elements, but the H1 pattern appeared anti-correlated with *telobox* density and with TRB1 peaks (Figure 5A and S13). To assess whether this apparent antagonism between H1 and TRB1 is a general property outside of *teloboxes*, we examined H1 occupancy over all TRB1 genome-binding sites and also found an inverse correlation between H1 and TRB1 on genes and TEs (Figure 5B). Consistent with the hypothesis that H1 prevents PRC2 recruitment over telomeric repeats, these observations hint at an antagonistic *cis*-enrichment of H1 and TRB1 over *telobox*-rich regions, including *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). We plotted the profiles of H1, TRB1, *telobox* motifs, and nucleosome occupancy over all WPNs. As expected, H1.2-GFP distribution matched the DNA linker regions while TRB1 tends to have a much broader distribution spanning 4-5 nucleosomes. The TRB1 profile nonetheless displayed a periodic pattern matching linker DNA coordinates, and, surprisingly, the *telobox* distribution itself sharply coincided with regions serving as linker DNA (Figure 5C). Considering this unexpected preferential distribution over linker DNA, we concluded that *telobox* motifs may play a functional role in nucleosome distribution and are subject to antagonistic binding by H1 and TRB1 (schematized in Figure 5D).

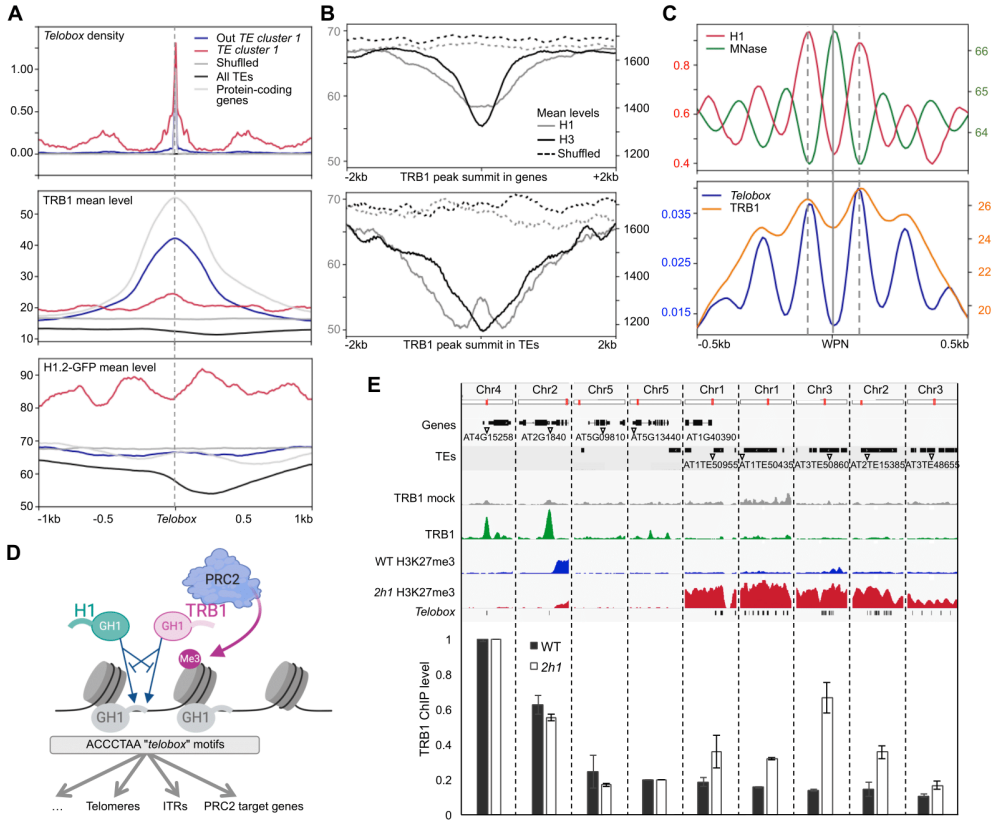



**Figure 5. H1 restricts the accessibility of TRB1 to *telobox* loci. A.** Global occupancy of TRB1 and H1 at all *telobox* positions in WT plants. TRB1 peak summit distribution expectedly correlates with the position of *teloboxes* located in protein-coding genes. In contrast, TRB1 is barely associated to the multiple *telobox* motifs of *TE cluster 1* elements, which are globally H1-rich. H1 and TRB1 patterns appear to be both influenced by *telobox* positioning, and 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 tested whether H1 depletion triggers ectopic TRB1 recruitment to *telobox*-containing *TE Cluster 1-2* elements 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 *teloboxes* 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 selected from *TE cluster 1* and three interspersed *ATREP18* or *LTR-Gypsy* elements selected from *TE cluster 2* that are located on other chromosomes. Corroborating the ChIP-seq analyses, no significant

TRB1 association was detected over these five loci in WT plants but TRB1 association was reproducibly detected in *2h1* samples (Figure 5E bottom panel). These findings shed light on a role for H1 in preventing TRB recruitment on *TE Cluster 1* and *2* elements, therefore providing a plausible mechanism for *de novo* H3K27me3 deposition on ITRs in the absence of H1.

**A role for H1 in telomere chromatin composition and sub-nuclear positioning**

Considering that telomeres display a continuum of perfect *telobox motifs* and a propensity to attract TRB proteins (Schrumpfová et al., 2014), we assessed whether these chromosome domains were, similar to ITRs, 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 anti-H3K27me3 and anti-H3 ChIP DNA to radioactively labeled concatenated telomeric probes (Adamusová et al., 2020). This led to the estimation that telomeres display an average ~4-fold H3K27me3 enrichment in *2h1* as compared to WT plants, independent of any detectable change in nucleosome occupancy (Figure 6A and S14).

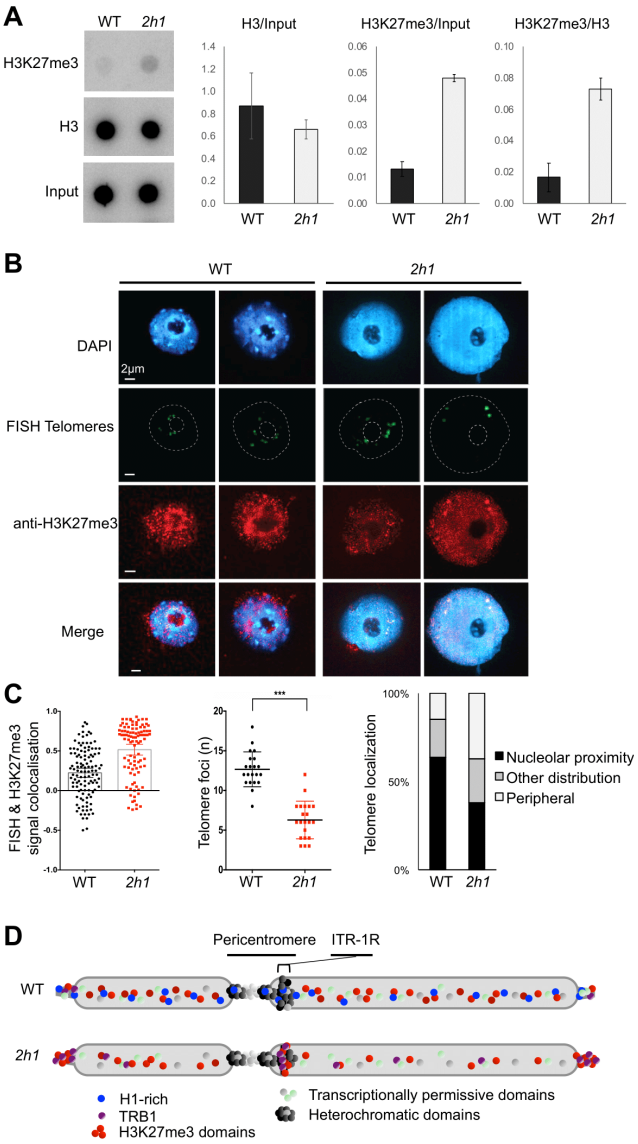



**Figure 6. H1 influences H3K27me3 enrichment and the sub-nuclear organization of telomeres. A.** Telomere ChIP analysis of H3K27me3 signal relative to nucleosome occupancy is enhanced in *2h1* plants. Anti-H3K27me3 ChIPs were followed by dot-blot hybridization with a radioactively labeled telomeric 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.

To assess whether H3K27me3 enrichment concerns a few telomeres or affects them generally, we explored the nuclear distribution of this histone mark by immunolabeling combined with telomere Fluorescence In Situ Hybridization (DNA FISH). Expectedly, in WT nuclei most visible telomeric foci were usually distributed around the nucleolus and colocalized with H3K27me3 signals. Consistent with our dot-blot analysis, H3K27me3 signal intensity at telomere foci was enhanced in the *2h1* line (Figure 6B-C). Moreover, 2-to-4 telomere FISH foci frequently presented outstandingly strong H3K27me3 labeling (Figure 6B). We did not ascertain whether some of these strong signals correspond to cross-hybridizing pericentromeric ITRs. Their frequent positioning near to the nuclear periphery coincides with the typical localization of pericentromeres and might point out to the latter hypothesis, albeit many of the telomeres were also abnormally distant from the nucleolus in *2h1* plants (Figure 6C).

In this analysis, we 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 from indirect topological alterations leading to their mislocalization in the nucleus. Collectively, we concluded that H1 does not only prevent accumulation of H3K27me3 over ITRs and telomeres but also influences the sub-nuclear organization of interphase chromosomes (Figure 6D).

**H1 promotes 3D chromatin packaging but attenuates the insulation of ITRs**

To gain a more detailed view of the defects in chromosome organization induced by *H1* loss-of-function and to investigate how telomeres and ITRs could be impacted, we employed *in situ* Hi-C on WT and *2h1* nuclei isolated from dissected cotyledons. The tissue homogeneity and high read coverage allowed us to reach a high 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 but fewer within the chromosomal arms (Figure S16A). Comparison of interaction frequencies as a function of genomic distance unveiled a tendency for fewer long-range interactions within chromosome arms and pericentromeric regions in *2h1* mutant nuclei (Figure 7A). Determination of interaction decay exponents (IDEs), which characterize chromatin packaging, confirmed a less steep decay with distance in *2h1* nuclei (Figure 7B). Both observations are consistent with a general trend for chromatin decompaction in the absence of H1. Accordingly, visualization of differential interaction frequencies between WT and *2h1* nuclei at a 100 kb resolution shed light on a clear decrease of intra-pericentromeric long-range interactions (i.e., blue squares surrounding the centromeres) and an increase association frequency of pericentromeric-regions with their respective chromosome arm (i.e., red crosses), which together reflects the loosening of pericentromeric heterochromatin compaction in *2h1* cotyledon nuclei (Choi et al., 2019; Rutowicz et al., 2019) (Figure 7C and S17).

We examined more precisely ITR-1R and 4L and observed that both chromatin regions form compacted structures resembling topologically associating domains (TADs) in WT plants (Figure 7D, S16A and S17A). Interestingly, intra-ITR interactions were strongly enhanced in *2h1* plants (Figure 7D, S16 and S17) indicating an enhanced insulation of the two ITR TAD-like structures in the absence of H1. Hence, the deletion of H1 triggered opposite trends of 3D chromatin organization defects between the two ITRs and their neighboring pericentromeric environments, the latter being characterized by extensive relaxation of heterochromatin.

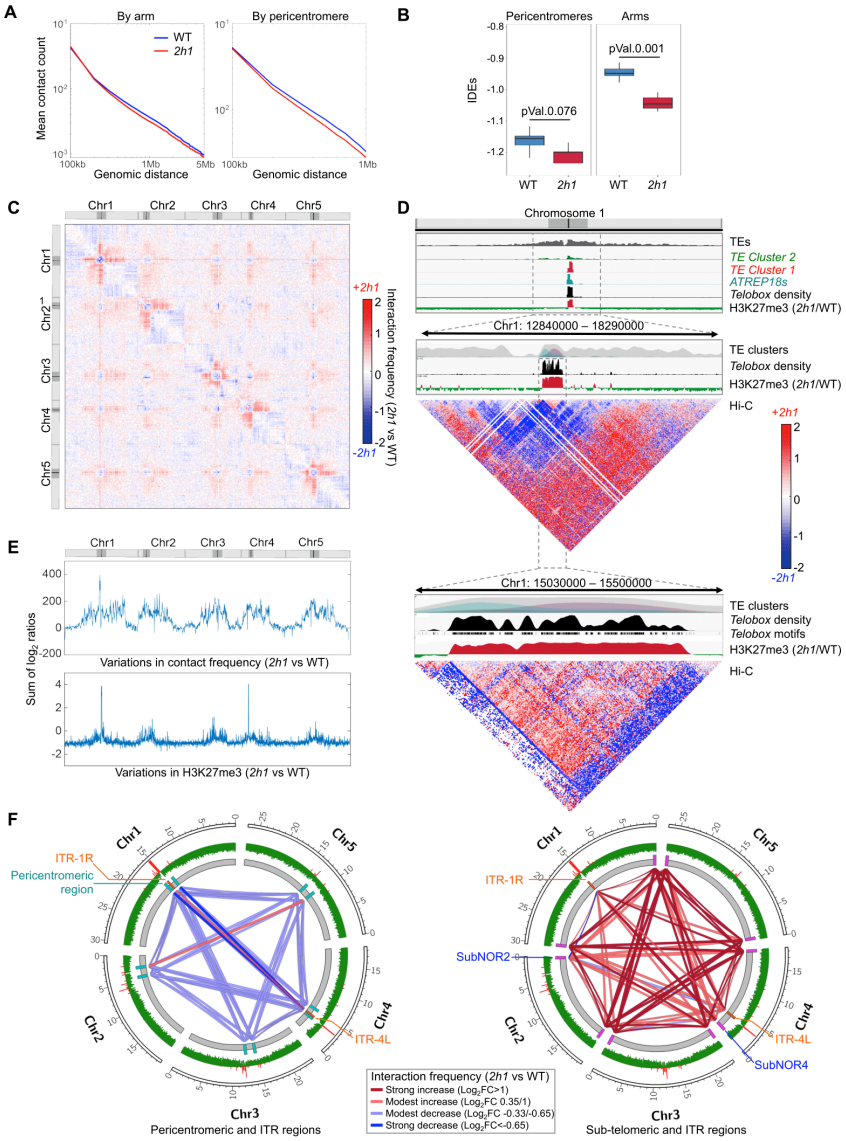



**Figure 7. H3K27me3 accumulation on ITR-1R and on the telomeres associates with higher-order topological defects.**
**A.** Mean contact count as a function of genomic distance for all chromosomes at a 100 kb resolution. **B.** Distribution of IDEs determined at a 100 kb resolution for chromosome arms and pericentromeric regions of WT and *2h1* nuclei. Median IDE values of chromosome arms and pericentromeres were determined as -0.95/-1.16 in WT and -1.05/-1.2 in *2h1* nuclei, respectively. P-values for pairwise Mann–Whitney U test comparisons between replicates are reported. **C.** Relative differences of interaction frequency between WT and *2h1* plants. The $Log_2$ values of interaction frequency of the five chromosomes in *2h1* versus WT are shown at a 100 kb resolution. Regions in red have more frequent contacts in *2h1* than in WT plants; regions in blue have an opposite behavior. The top and side tracks show the positions of centromeres (black) and pericentromeric domains (grey). **D.** In *2h1* mutant plants, a general reduction of intra-pericentromeric interaction frequencies is contrasted by enhancement of TAD formation matching the H3K27me3-enriched ITR-1R 355 kb block. Top panel, location of ITR-1R in chromosome 1. Middle panel, magnification of the region surrounding chromosome 1 pericentromeres at a 10 kb resolution. Bottom panel, magnification of the pericentromere-imbedded ITR-1R at a 2 kb resolution. **E.** Comparison of Hi-C and H3K27me3 genome-wide variations 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 levels using 10kb bins. **F.** Variations of inter-chromosomal interaction frequencies between telomere-proximal, pericentromeric and ITR-1R and 4L regions. Except for the SubNOR2 and SubNOR4 regions that displayed specific alteration patterns, interactions between telomere-proximal regions and with the two ITRs tend to be more frequent in the absence of H1 (see also Figures S16C and S17C). In contrast, interactions are less frequent between pericentromeres or between ITR-1R/4L with pericentromeric regions of other chromosomes. Yellow boxes, ITR regions. External green/red track, H3K27me3 variations in *2h1* versus WT plants (Log2 ratio). Magenta boxes, telomere proximal regions and SubNOR2 or 4. Hi-C data were analyzed using three independent biological replicates.

Magnification of the ITR-1R region at a 2 kb resolution revealed a remarkably sharp correspondence between these TAD boundaries and H3K27me3-enrichment in *2h1* plants (Figure 7D bottom panel). To further assess the relationships between topological and H3K27me3 defects in *2h1* plants, we plotted all differences in H3K27me3 profiles over the chromosomes (sum of log2 ratios) and compared them to a similar representation of the Hi-C data (Figure 7E). The comparison sheds light on ITR-1R as a major H3K27me3-enriched genome locus with prominent differences in chromatin interactions in *2h1* chromatin.

**H1 impacts long-distance interactions between chromosome ends**

Considering that H3K27me3-enriched telomeres display altered sub-nuclear organization in *2h1* mutant plants (Figure 6B-C), we examined long-distance interactions among the corresponding chromosome regions in the Hi-C matrices. 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 sub-telomeric regions. In this analysis, we also considered an internal 100-kb region of each pericentromeric region, the ITR-1R and 4L, and several 100-kb regions randomly chosen in distal chromosomal arms. Possibly reflecting the capacity of different centromeres to aggregate within chromocenters, in WT plants we observed that pericentromeric regions frequently interact with each other as compared to chromosomal arm domains (Figure S17A-B). Similarly, in agreement with the preferential localization of centromeres at the nuclear periphery and of the telomeres around the nucleolus in the nuclear interior (reviewed in Pontvianne and Grob, 2020; Santos et al., 2020), most of the telomere-proximal regions frequently interacted with each other through long-range interactions, but less with ITR-1R and 4L (Figure S17A). One exception was the 100-kb region (referred here as the SubNOR4 region) that is separated from chromosome 4 telomere by the ~4 Mb Nucleolar Organizing Region 4 (*NOR4*) (Copenhaver and Pikaard, 1996), which encodes highly transcribed 45S ribosomal RNA genes (Mohannath et al., 2016) and tends to be more isolated than all other telomere proximal regions (Figure S17A).

In *2h1* nuclei, we observed that, with the exception of the SubNOR2 and SubNOR4 regions that displayed specific defects, the frequencies of interaction between the different sub-telomeric regions were increased (Figure 7F and S17). 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 4L did not show such a clear tendency, apart from a decreased frequency of association with pericentromeric regions as detailed above, and a slight increase between them and sub-telomeric regions (Figure 7F and S17C). In brief, in contrast to its positive impact on intra- and inter-pericentromere associations, H1 appears to dampen long-distance associations between telomeres in addition to preventing H3K27me3 enrichment on these loci.

## Discussion

We report that *Arabidopsis* H1 is highly enriched at PRC2 target genes where it contributes to efficient H3K27me3 marking and to low chromatin accessibility. Contrasting with this general tendency, we identified an opposite role of H1 in limiting H3K27me3 deposition over interstitial telomeric repeats,

telomeres, and a few genes frequently displaying TE-like features or containing a *telobox* sequence motif in their promoter region. Thus, H1 has a differential effect on H3K27me3 levels over thousands of protein-coding genes on the one hand and over loci characterized by repeated telomeric motifs on the other hand. This antagonistic behavior provides an explanation for our former observation that most of H3K27me3 nuclear signals are low in *H1* loss-of-function plants whilst, intriguingly, a few foci of undetermined nature remain enriched (Rutowicz et al., 2019). The large scale on which these antagonistic patterns are observed sheds light on the existence of strong mechanistic links between H1 and *Polycomb*-based regulation in *Arabidopsis*, two main factors in the instruction of DNA accessibility.

**Promoting H3K27me3 enrichment on genes: an evolutionarily conserved function of H1**

A majority (57%) of the H3K27me3-marked genes were hypo-methylated in *2h1* plants, yet, most of the H3K27me3 peaks are not completely erased. This pattern suggests that H1 has a general influence on H3K27me3 deposition/maintenance or spreading but is not mandatory for the nucleation of this mark. Although we have not been able to identify any sequence specificity for H1-mediated H3K27me3 enrichment over most protein-coding genes, the dual influence of H1 on chromatin status and on transcriptomic patterns underlie its implication in selective mechanisms. While H1 depletion typically resulted in a global increase of gene chromatin accessibility, its impact on expression was apparently more related to variations in H3K27me3 marking. Hence, consistent with the functional categories of the misregulated genes, and with subtle phenotypes of *H1*-knockout plants reminiscent of PRC2 loss-of-function plants (Rutowicz et al., 2019), part of the defects in gene expression observed after H1 depletion might result from indirect consequences through PRC2 mis-function. These properties might represent an evolutionarily conserved function of H1, as suggested by recent findings in mouse where the depletion of H1 variants triggers widespread H3K27me3 loss and the misregulation of PRC2-regulated genes, phenocopying loss of EZH2 (Willcockson et al., 2020; Yusufova et al., 2020).

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

Considering their capacity to recruit the somatic PRC2 methyltransferases CURLY-LEAF (CLF) and SWINGER (SWN) to *telobox*-containing genes (Zhou et al., 2018), TRBs represent excellent candidates for the sequence-specific regulation of H3K27me3 on ITRs and possibly also telomeres. Waiting for an assessment of the relative affinity of H1 and TRB1 for *telobox* elements in a chromatin context, several observations support a model in which H1 antagonizes PRC2 activity on telomeric repeats through competitive binding to DNA with TRB proteins. Firstly, on a genome-wide scale, H1 and TRB1 occupancies are negatively correlated. Secondly, analysis of nucleosome positioning showed that *telobox* motifs are preferentially situated in linker DNA where TRB1 association is also pronounced; hence competition with H1 might occur on linker DNA. Last, supporting this model, we showed that TRB1 ectopically invades ITRs and other *telobox*-rich elements gaining H3K27me3 in *2h1* plants. This indicates that, in wild-type plants, elevated H1 incorporation on these loci limits TRB1 enrichment and/or accessibility despite the presence of repeated *telobox* motifs for which the Myb domain of TRB1 has strong affinity (Mozgová et al., 2008; Schrumpfová et al., 2016). Given its highly dynamic association with chromatin (Dvořáčková et al., 2010), ectopic enrichment of TRB1 on ITR regions in *2h1* plants may

therefore simply result from an increased accessibility of *telobox* sequences otherwise occupied by H1 (Figures 5D).

Our model echoes the recent report that *Arabidopsis* GH1-containing High Mobility Group A1 (GH1-HMGA1) is present at telomeres and minimizes H1 incorporation into chromatin (Charbonnel et al., 2018; Kotliński et al., 2017). Hence, in line with mammalian HMGs and H1 antagonistically binding to DNA (Catez et al., 2004; Krishnakumar et al., 2008), *Arabidopsis* H1 chromatin incorporation might act in competition with several GH1-domain-containing proteins, such as TRB1 and GH1-HMGA1. In *Arabidopsis*, the 15 GH1-domain-containing proteins (Charbonnel et al., 2018; Kotliński et al., 2017) may allow playing this H1 theme using multiple combinations of antagonistic or cooperative interactions to regulate chromatin.

Future studies will determine whether H3K27me3-enrichment on telomeric repeats directly relies on PRC2 recruitment by TRB proteins, as recently shown using *telobox*-containing promoter reporter lines (Zhou et al., 2018), or whether other chromatin modifiers influencing H3K27me3 are implicated. At this stage the latter possibility cannot be discarded as, for example, the *Arabidopsis* PRC1 subunit LIKE-HETEROCHROMATIN 1 (LHP1) acting as a chromatin reader of H3K27me3 (Turck et al., 2007), prevents TRB1 enrichment on PRC2 target genes displaying *telobox* motifs (Zhou el al., 2016). The outstanding pattern of *telobox* positioning in linker DNA also suggests a capacity of this sequence motif to influence chromatin organization, possibly by repelling nucleosomes away from *telobox* motifs.

**A new role for H1 on telomeric chromatin structure**

Owing to their repetitive nature (Fojtová and Fajkus, 2014; Majerová et al., 2014; Vaquero-Sedas et al., 2011, 2012; Vega-Vaquero et al., 2016), the chromatin composition and organization of plant telomeres has long remained enigmatic (Achrem et al., 2020; Dvořáčková et al., 2015). ChIP dot-blot analyses indicated a dominance of H3K9me2 over H3K27me3 histone marks but some of the *Arabidopsis* telomere regions also display euchromatic H3K4me2/3 marks (Adamusová et al., 2020; Grafi et al., 2007; Vaquero-Sedas et al., 2011). Here, combining ChIP-seq, ChIP dot-blot with telomeric probes and *in situ* immunolocalization showed that H1 moderates by 2-to-4 fold the accumulation of this mark on telomeres. However, our analyses did not allow assessing the precise distribution of HK27me3 enrichment along each telomere, especially considering their mosaic chromatin status. *Arabidopsis* telomeres are indeed thought to comprise chromatin segments with distinct nucleosome repeat length (NRL), for instance one of an average NRL of 150 bp (Ascenzi and Gantt, 1999), which is much shorter than the 189 bp estimated for H1-rich TEs (Choi et al., 2020). Such a small DNA linker size, i.e., 5 bp, is seemingly incompatible with H1 incorporation into chromatin, as H1 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 it is 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 some telomere segments would exist in an H1-free state and display a columnar structure in which nucleosome arrays are stabilized be stacking interactions mediated by the histone octamers themselves (Fajkus and Trifonov, 2001). In conclusion, the existence of distinct chromatin states at *Arabidopsis* telomeres needs to be explored in more detail to establish whether the repressive influence of H1 on PRC2 activity is a general property of telomeres or rather impacts specific segments.

**H1 has a profound influence on the *Arabidopsis* 3D genome topology**

Using Hi-C we identified a reduced frequency of chromatin interactions within and among the pericentromeres in *2h1* nuclei. This is a typical feature of *Arabidopsis* mutants affecting chromocenter formation (Feng et al., 2014; Grob et al., 2014; Moissiard et al., 2012) or when chromocenters get decompacted in response to environmental stress (Sun et al., 2020). These analyses 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). They also shed light on a complex picture in which ITR-1R and 4L embedded within the pericentromeres of chromosomes 1 and 4 escape the surrounding relaxation of heterochromatin induced by H1 depletion and organize themselves as TAD-like structures. H3K27me3 *cis*-enrichment resulting from H1 depletion might underlie the maintenance of compacted and very poorly accessible ITR chromatin while neighboring heterochromatic regions tend to become more accessible.

We also observed that H1 depletion leads to a reduction in the proportion of telomeric foci located near the nucleolus and in their total number. Using Hi-C, we could attribute this apparent lack of telomere spatial individualization to more frequent inter-chromosomal interactions between telomere proximal regions, which were used as a proxy for telomeres in the Hi-C analysis. As the preferential positioning of telomeres around the nucleolus and of the centromeres near to the nuclear periphery are important organizing principles of *Arabidopsis* chromosome territories (recently reviewed in Pontvianne and Grob, 2020; Santos et al., 2020), H1 therefore appears to be a crucial determinant of *Arabidopsis* interphase nuclear organization.

While both PRC1 and PRC2 participate in defining *Arabidopsis* genome topology (Feng et al., 2014; Veluchamy et al., 2016), H3K27me3 is favored among long-distance interacting gene promoters (Liu et al., 2016). This led to the proposal that, as in animals, this mark could contribute to shape chromosomal organization in *Arabidopsis*, 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 telomeres, pericentromeres and ITR regions. In mammals, H1 depletion not only triggers higher-order changes in chromatin compartmentation (Willcockson et al., 2020; Yusufova et al., 2020), but also extensive topological changes of gene-rich and transcribed regions (Geeven et al., 2015). Future studies will help to determine to which extend the impact of H1 on the H3K27me3 landscape contributes to define *Arabidopsis* genome topology.

**H1 as a modulator of H3K27me3 epigenomic homeostasis**

With thousands of perfect and degenerated telomeric motifs altogether covering ~430 kb, ITR-1R and 4L represent at least twice the length of all *Arabidopsis* telomeres combined, which span 2 to 5 kb at the end of each chromosome (Fitzgerald et al., 1999; Richards and Ausubel, 1988). Therefore, as illustrated in Figure 6D, the cumulated presence of thousands *teloboxes* in ITR-1R and 4L forms an immense reservoir of PRC2 targets. Our observations show that H1 serves as a safeguard to avoid the formation of gigantic H3K27me3-rich blocks in both pericentromeric ITRs and telomeres, which would be on a scale potentially tethering many PRC2 complexes away from protein-coding genes. In *Neurospora crassa*, artificial introduction of an array of $(TTAGGG)_{17}$ telomere repeats at interstitial sites triggers the formation of a large block of H3K27me2/3-rich chromatin (Jamieson et al., 2018). This example and our

findings illustrate the intrinsic attractiveness of telomeric motifs for H3K27me3 deposition in multiple systems. Our study further suggests the implication of H1 in balancing PRC2 activity between protein-coding genes and telomeric repeats in plants, hence potential acting as a modulator of the epigenome's homeostasis.

## Methods

### Plant lines and growth conditions

All plants used in this study correspond to *A. thaliana* Col-*0.* The *h1.1 h1.2* (*2h1*) *Arabidopsis* mutant line and the transgenic *pH1.2::H1.2-GFP* line (Rutowicz et al., 2015) were kindly provided by Dr. Kinga Rutowicz. 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 $\mu mol.m^{-2}.s^{-1}$) for 5 days unless otherwise stated. Cotyledons, when used, were manually dissected under a stereomicroscope.

### Immuno-FISH

After fixation in 4% paraformaldehyde in 1X PME, 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. 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% PFA in 1X PME was added and slides were dried. 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 FACS according to their ploidy levels using a MoFlo Astrios EQ Cell Sorter (Beckman Culture) in PuraFlow sheath fluid (Beckman Coulter) at 25 psi (pounds per square inch), with a 100-micron nozzle. We performed sorting with ~43 kHz drop drive frequency, plates voltage of 4000-4500 V and an amplitude of 30-50 V. Sorting was performed in purity mode. 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) with downscaling using seedlings crosslinked in 10 mM potassium phosphate pH 7.0, 50 mM NaCl, 0.1 M sucrose with 4 % (v/v) formaldehyde. Crosslinking was stopped by transferring seedlings to 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 2% Triton X-100. DpnII (200 u) was added to each sample for over-night digestion at 37 °C. dATP, dTTP, dGTP, biotinylated dCTP and 12 µl DNA Polymerase I (Large Klenow fragment) were added before incubation for 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 to reach a final volume of 700ul. Samples were incubated for 4h at 16°C with constant shaking at 300rpm. After over-night reverse crosslinking 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 1.6X ratio. DNA was fragmented using a Covaris M220 sonicator (peak power 75W, duty factor 20, cycles per burst 200, duration 150 s). 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 1.85X). Libraries were analyzed using a Qubit fluorometer (Thermofisher) and a TAPE Station (Agilent) before sequencing in a 75 bp PE mode using a DNB-seq platform at the Beijing Genomics Institute (BGI Group; Honk Kong). Mapping of Hi-C reads was performed using the Hi-C Pro pipeline (Servant et al., 2015) with default pipeline parameters and merging data from three biological replicates at the end of the pipeline. Data were in visualized using the Juicebox toolsuite (Durand et al., 2016) and represented in $Log_{10}$ scale after SCN normalization (Cournac et al., 2012) with Boost-HiC (Carron et al., 2019) setting alpha parameter to 0.2. In Figure S17, we normalized the sequencing depth in each sample and scored the number of reads in each combination of genomic regions using HOMER (Heinz et al., 2010). Read counts

were further normalized for the bin size and the median value between the three biological replicates was reported.

**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 to identify differentially 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 by $10^6$. Mean TPM values between two biological replicates were used for subsequent analyses. To draw metagene plots, genes were grouped into expressed or not and expressed genes split into 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 subjected to 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).

**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 0.1X 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 a 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).

**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 (Illumina NextSeq 500, 1x50bp) of all input and IP samples by Fasteris (Geneva, Switzerland).

**ChIP-seq and MNase-seq bioinformatics**

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 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 and TEs using bedtools v2.29.2 multicov and converted them in median counts per million, 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-plot 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 the R statistical environment. All scripts used will be made publicly available. Shuffled controls, 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 (Lyons 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.

**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 identity score were calculated using Ident and Sim tool (https://www.bioinformatics.org/sms2/ident_sim.html).

**Data and materials availability**

This study did not generate new unique reagents. All public genomic data used in this study are listed in Additional file 4. Additional files will be made available upon request to the corresponding authors.

**Graphical abstracts** were created using Biorender.com

**Funding**

FB benefitted from grants of the Agence Nationale de la Recherche projects ANR-10-LABX-54, ANR-18-CE13-0004-01, ANR-17-CE12-0026-02 at IBENS and ANR-11-EQPX-0029, ANR-10-INBS-04; ANR-11-IDEX-0003-02 at the Imagerie-Gif core facility. Collaborative work between FB and CeB was supported by a research grant from the Velux Foundation (Switzerland) and the Ricola 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's 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 and ChIP-RX experiments; GT and MB generated ATAC-seq datasets; KA and MF performed telomere dot-blots; MB contributed to FACS nucleus sorting; 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, UG, PPS, JF, and SG supervised research, discussed the results and edited the manuscript. FB and ClB jointly wrote the paper and directed the study.

## Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Fredy Barneche (barneche@bio.ens.psl.eu).


## Acknowledgements

The authors are grateful to Erwann Cailleux (IBENS, Paris, France) and David Latrasse (IPS2, Orsay, France) for technical guidance with ATAC-seq; Nicolas Valentin (I2BC, Gif, France) for assistance with FACS; 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.

Choi, J., Lyons, D.B., Kim, M.Y., Moore, J.D., and Zilberman, D. (2020). DNA Methylation and Histone H1 Jointly Repress Transposable Elements and Aberrant Intragenic Transcripts. Molecular Cell *77*, 310-323.e7.

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., Královics, 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*, 46.

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 Harbor Perspect. Biol. *6*, 1–26.

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.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Molecular Cell *38*, 576–589.

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.

Mohannath, G., Pontvianne, F., and Pikaard, C.S. (2016). Selective nucleolus organizer inactivation in Arabidopsis is a chromosome position-effect phenomenon. Proceedings of the National Academy of Sciences of the United States of America *113*, 13426.

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.

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.

Willcockson, M.A., Healton, S.E., Weiss, C.N., Bartholdy, B.A., Botbol, Y., Mishra, L.N., Sidhwani, D.S., Wilson, T.J., Pinto, H.B., Maron, M.I., et al. (2020). H1 histones control the epigenetic landscape by local chromatin compaction. Nature.

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.

Yusufova, N., Kloetgen, A., Teater, M., Osunsade, A., Camarillo, J.M., Chin, C.R., Doane, A.S., Venters, B.J., Portillo-Ledesma, S., Conway, J., et al. (2020). Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. Nature.

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, 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

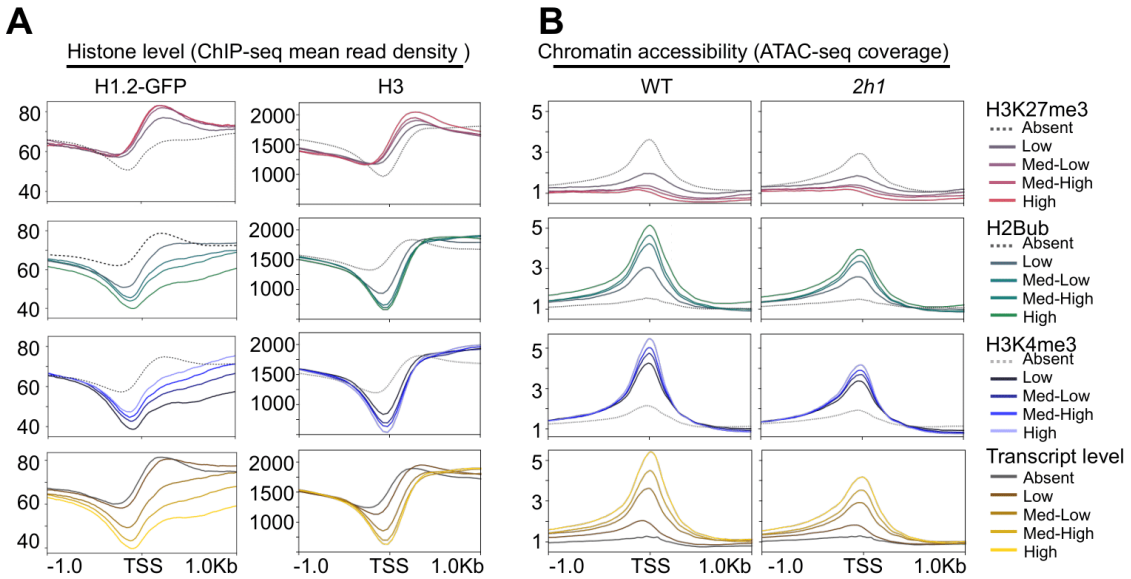



**Figure S1. H1.2 distribution and DNA accessibility on distinct chromatin states. A.** H1.2-GFP and H3 levels over the TSS of genes marked by different histone modifications characteristic of PRC2-based repression (H3K27me3; n=7542), of transcription initiation (H3K4me3; n=18735), transcription elongation (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

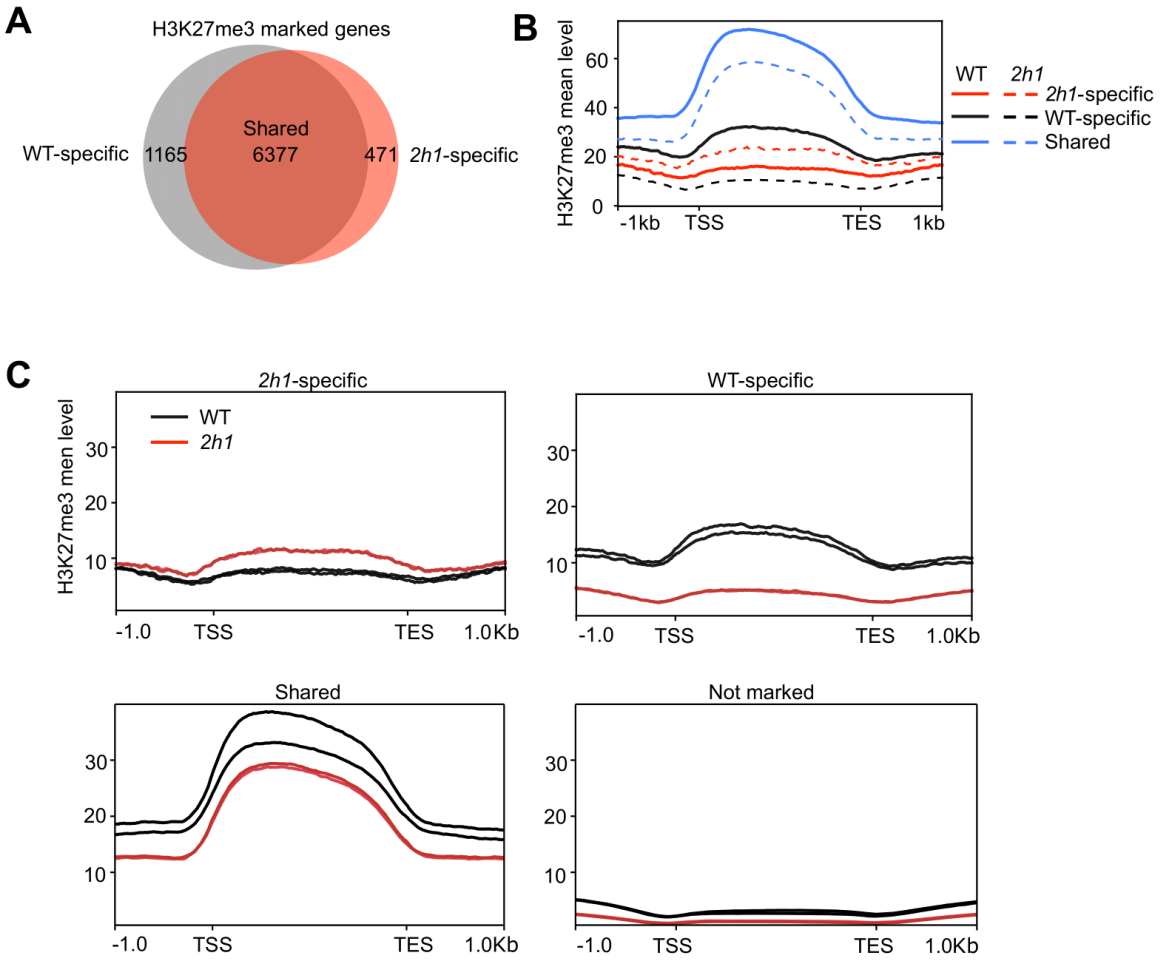



**Figure S2. Comparison of the gene repertoire significantly marked by H3K27me3 in WT and *2h1* plants. A.** Number of H3K27me3-marked genes in WT and *2h1* seedlings. **B.** H3K27me3 mean levels over the sets of genes acquiring (*2h1*-specific) or losing (WT-specific) H3K27me3 marking in *2h1* plants.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

Figure S3

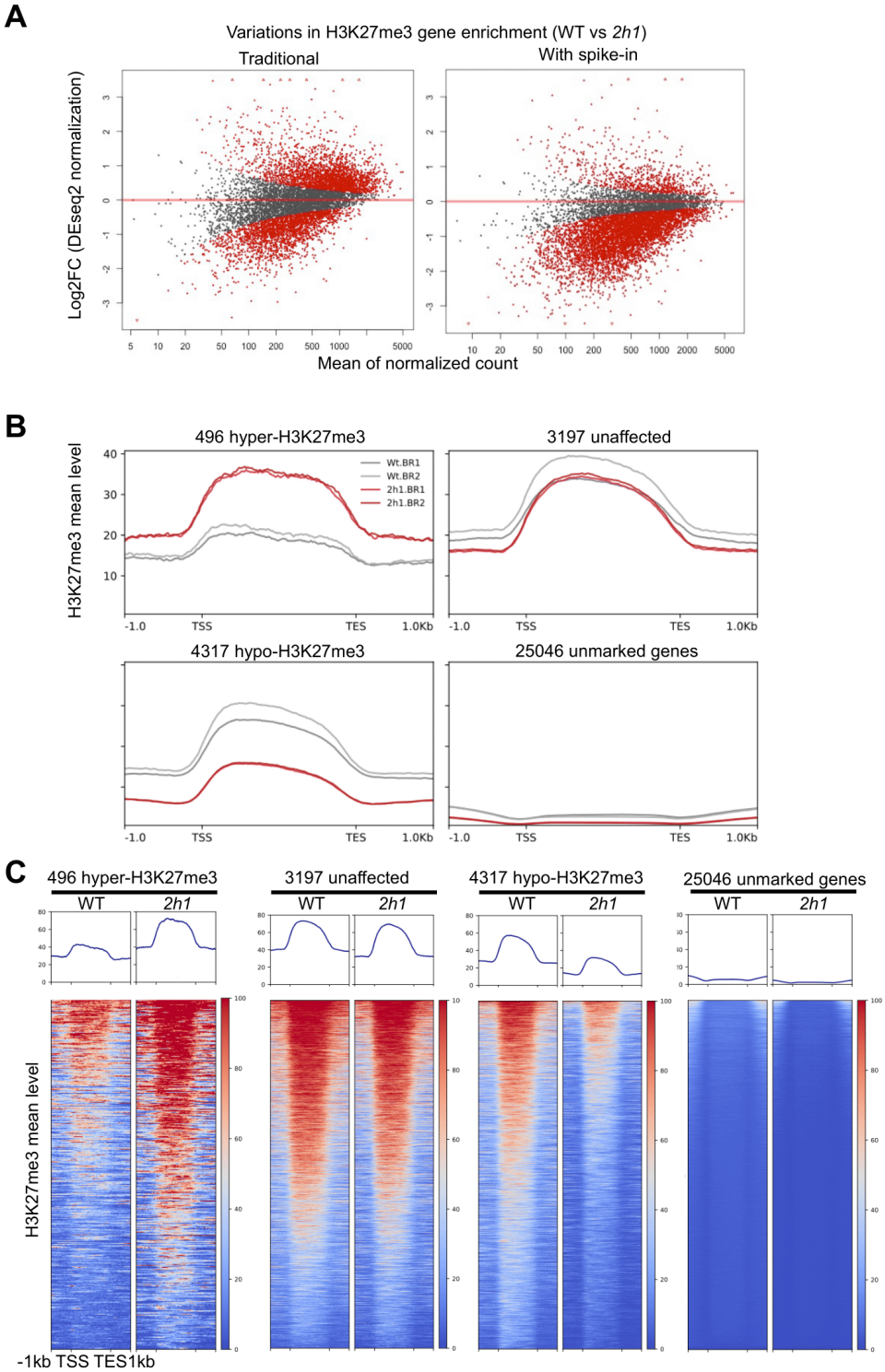



**Figure S3. Detailed analysis of H3K27me3 ChIP-Rx differential analyses. A.** Comparison of the DEseq2 results of two independent biological replicates (BR) using either a spike-in normalization factor or traditional Deseq2 normalization. **B.** Variability of H3K27me3 mean levels on differentially 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

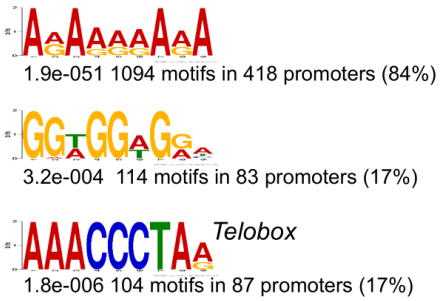



**Figure S4. Sequence motifs over-represented in the 496 H3K27me3 genes hyper-marked by H3K27me3 in *2h1* plants (E-value < 1e-220).** E-values were calculated against random sequences. The two first motifs could not be matched to any previously known regulatory motif while the 3rd identified motif corresponds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

Figure S5

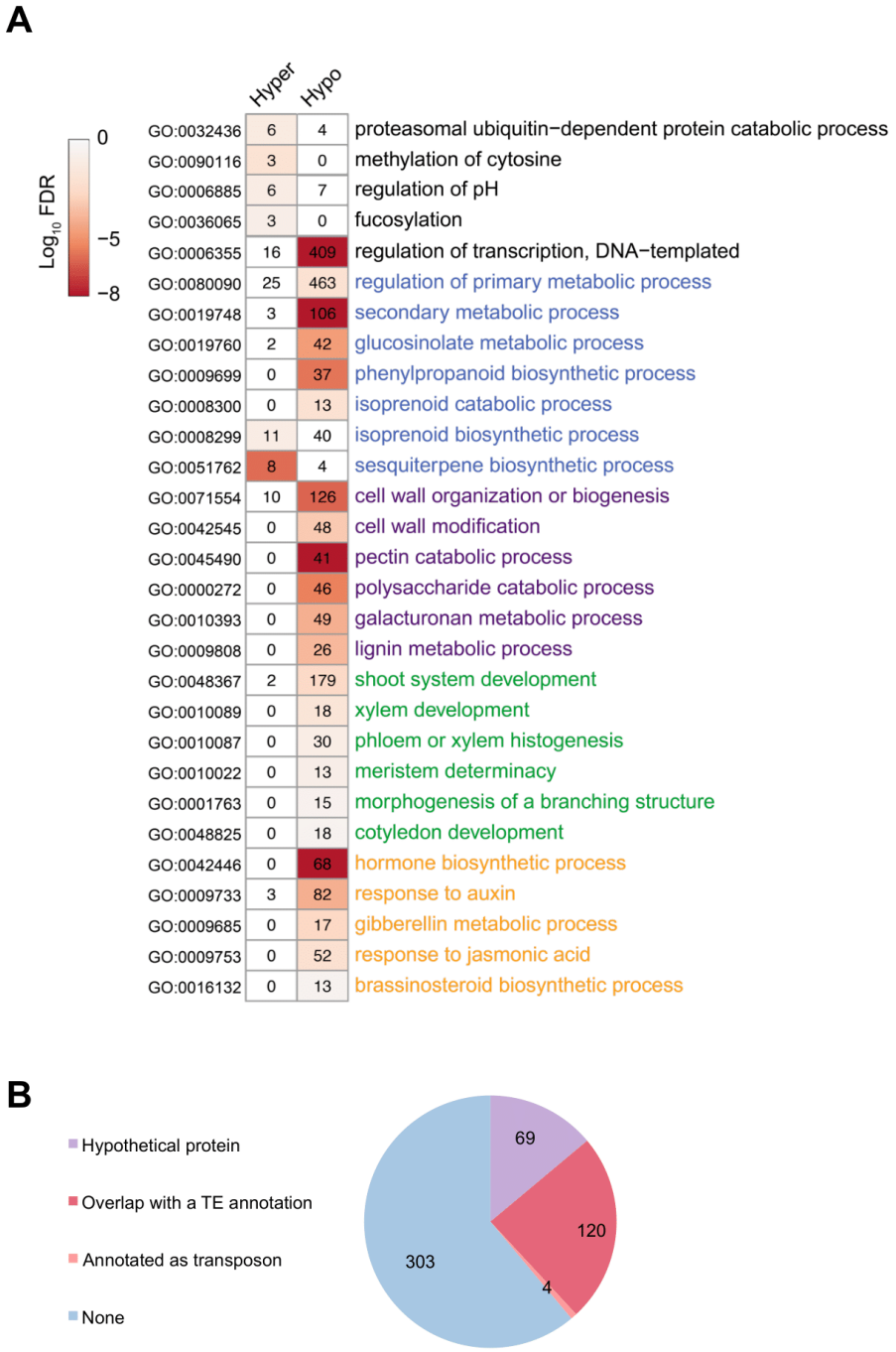



**Figure S5. A. Gene ontology analysis of H3K27me3 differentially marked genes in *2h1* plants. A.** Association to a significantly over-represented gene function is denoted as a heatmap of false discovery rate (FDR). N=4317 hypo-marked genes; 496 hyper-marked genes. **B.** Number of genes among the 496 hyper-marked genes that either overlap an annotated TE, are annotated as transposons, or are annotated as hypothetical proteins (See Additional file 1 for more details).

Figure S6

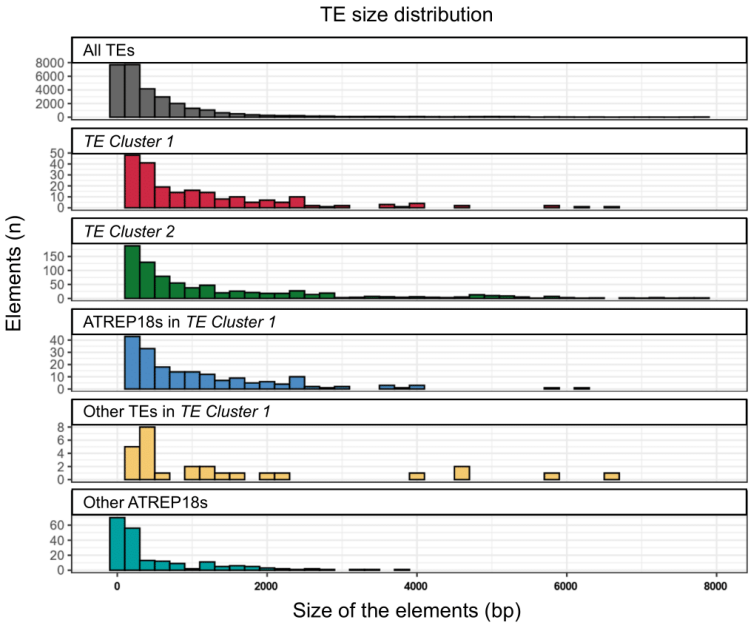



**Figure S6. Many *TE Cluster 1-2* elements tend to be longer than average TEs.** Size distribution of TEs and repeats belonging to the repertoires analyzed in Figure 3.

Figure S7

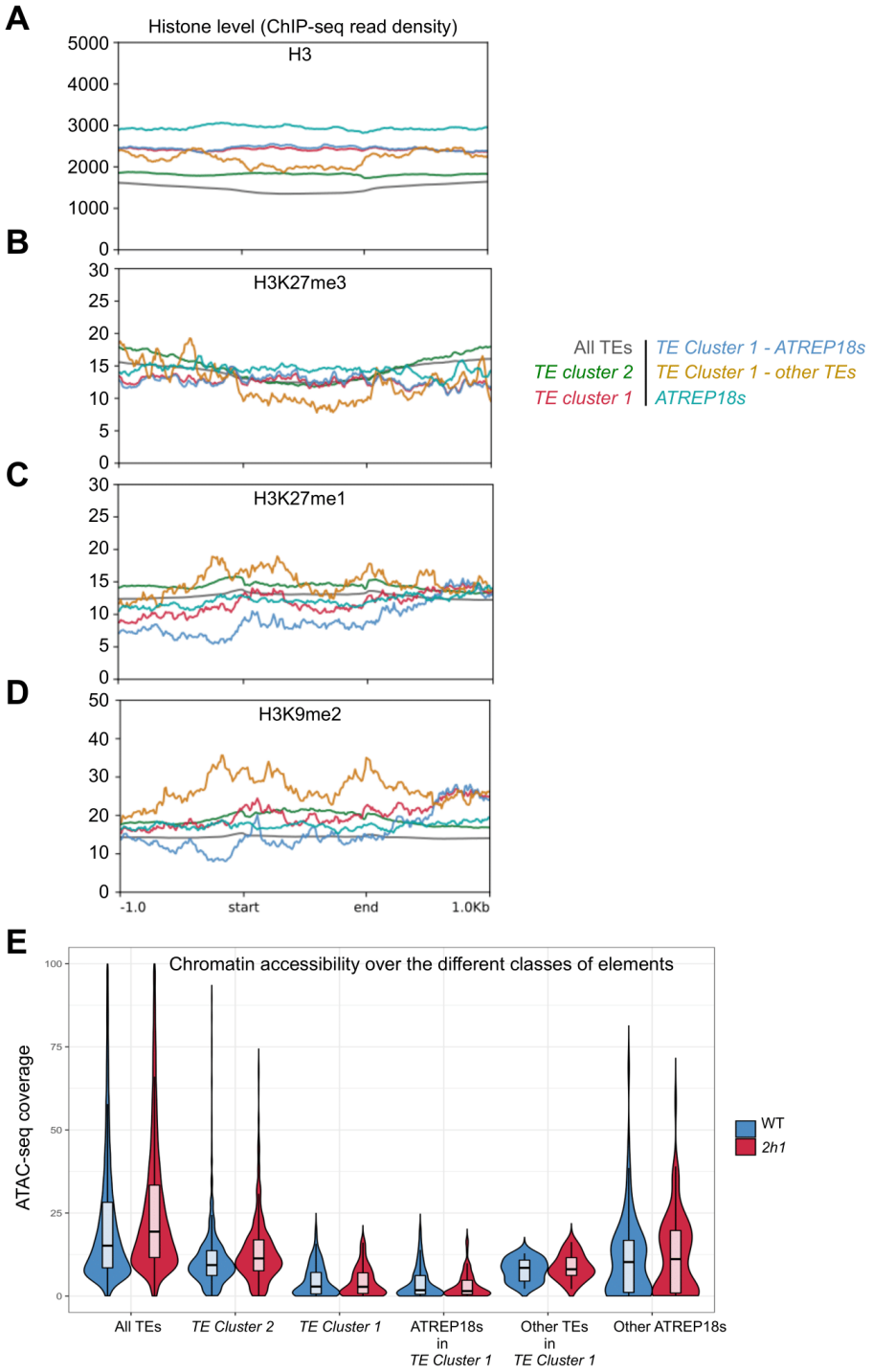



**Figure S7. Profiling of representative histone marks on *TE clusters 1-2* and on *ATREP18* elements. A.** H3 profiling indicates that *ATREP18s* and *TE Cluster1-ATREP18* elements display elevated nucleosome occupancy. This is consistent with the weak accessibility of these elements determined by ATAC-seq analyses. **B.** Comparison of H3K27me3 profiles indicates that all TE types investigated in this study, including *ATREP18* elements, display similarly low H3K27me3 levels in WT plants. **C.** Profiling of the heterochromatic H3K27me1 mark does not allow identifying a specific predisposition of *TE Cluster 1-2* or of *ATREP18* elements for H3K27me1 being used as a substrate for H3K27me3 enrichment. **D.** Profiling of the H3K9me2 mark on *TE Cluster 1-2*. All profiles have been generated in this study except H3K27me1 and H3K9me2 (Additional file 4). **E.** ATAC-seq mean read coverage along TEs and repeats from the indicated groups.

Figure S8

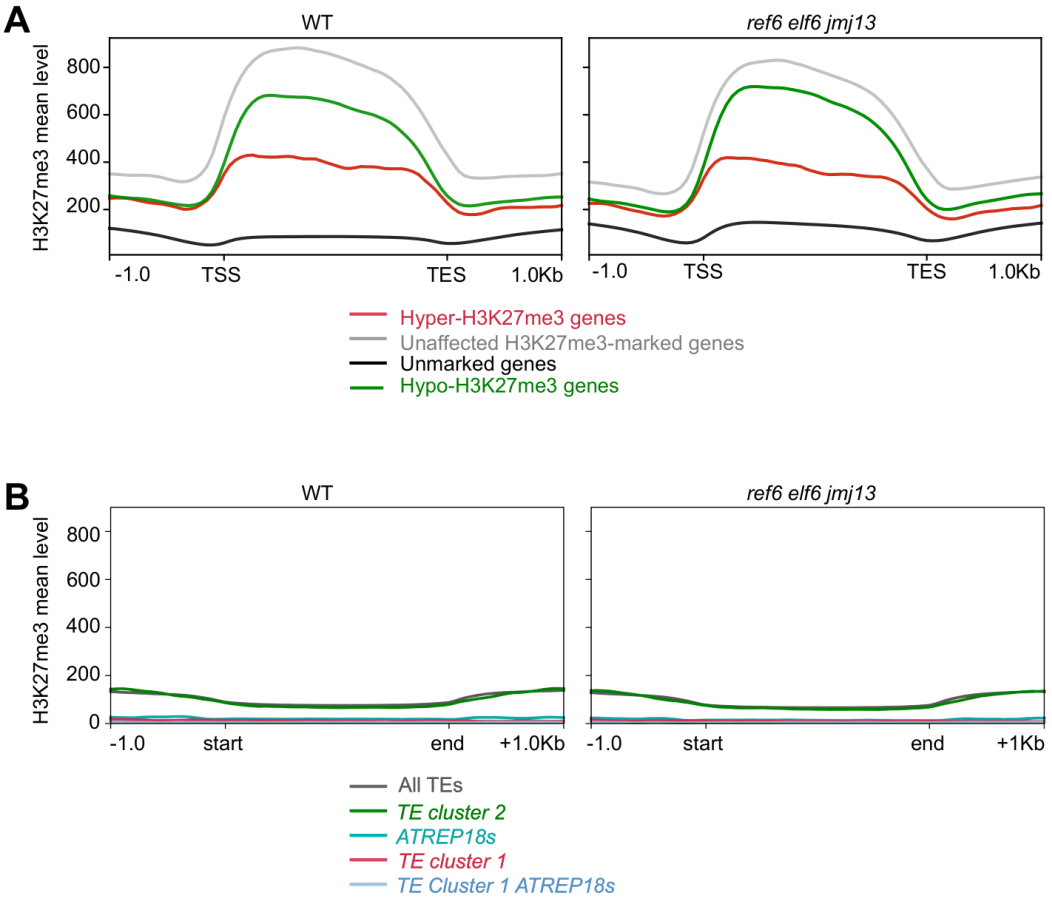
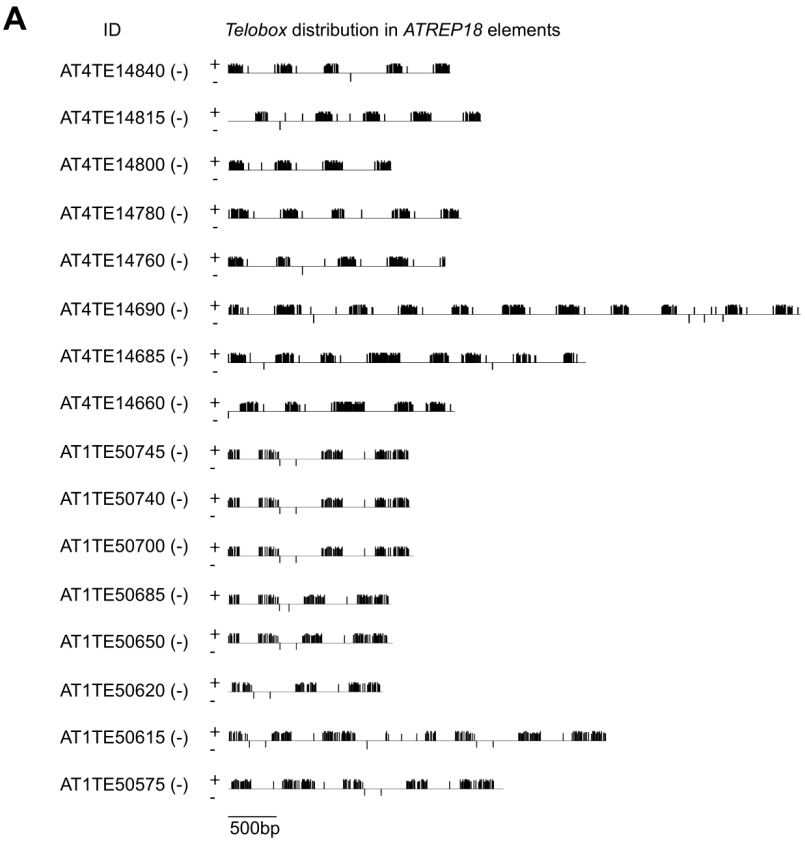



**Figure S8. H3K27me3 profiles of the indicated gene (A) and TE (B) sets in WT and *ref6 elf6 jmj13* triple mutant plants impaired in H3K27me3 demethylation.** H3K27me3 data were obtained from (Yan et al., 2018).

Figure S9



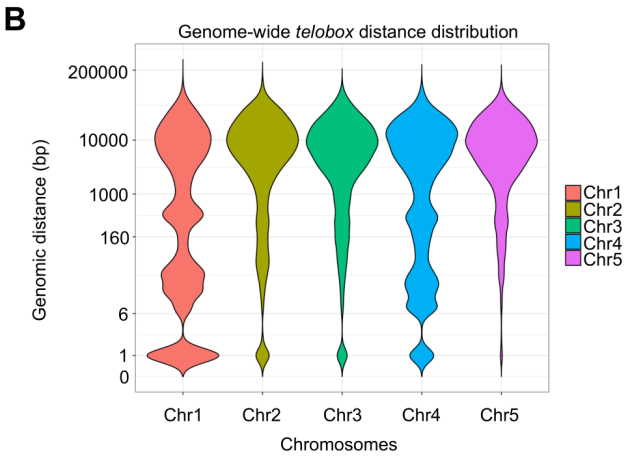




**Figure S9. *Telobox* motifs are organized as small clusters in *ATREP18* elements of chromosomes 1 and 4 interstitial telomeric regions. A.** *ATREP18* elements in *TE Cluster 1* display a particular pattern, with multiple clusters of *telobox* motifs. **B.** Genomic distances between all perfect *telobox*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

Figure S10

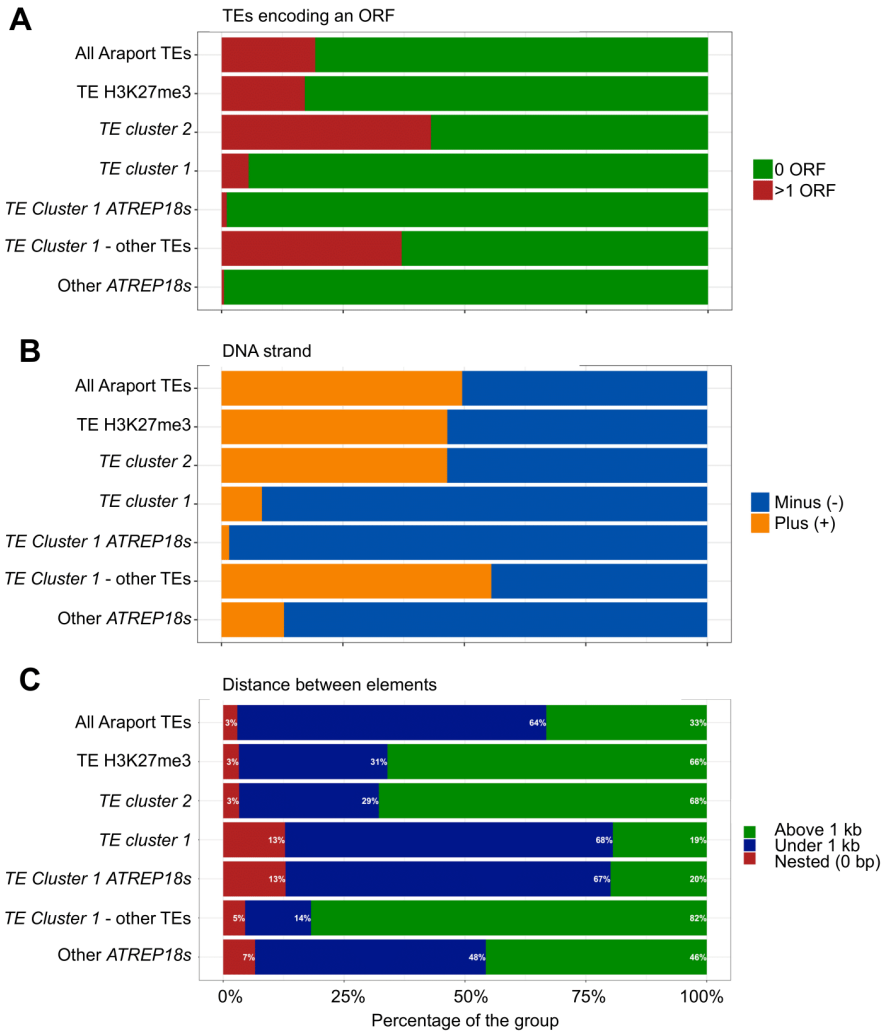



**Figure S10. Sequence features of *TE cluster 1-2* elements. A.** Frequency of open reading regions (ORFs) within the indicated sets of TEs. *TE Cluster 2*, which is made of a large variety of TE super-families dispersed in the genome, more frequently encodes ORFs compared to the ensemble of all TEs. 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

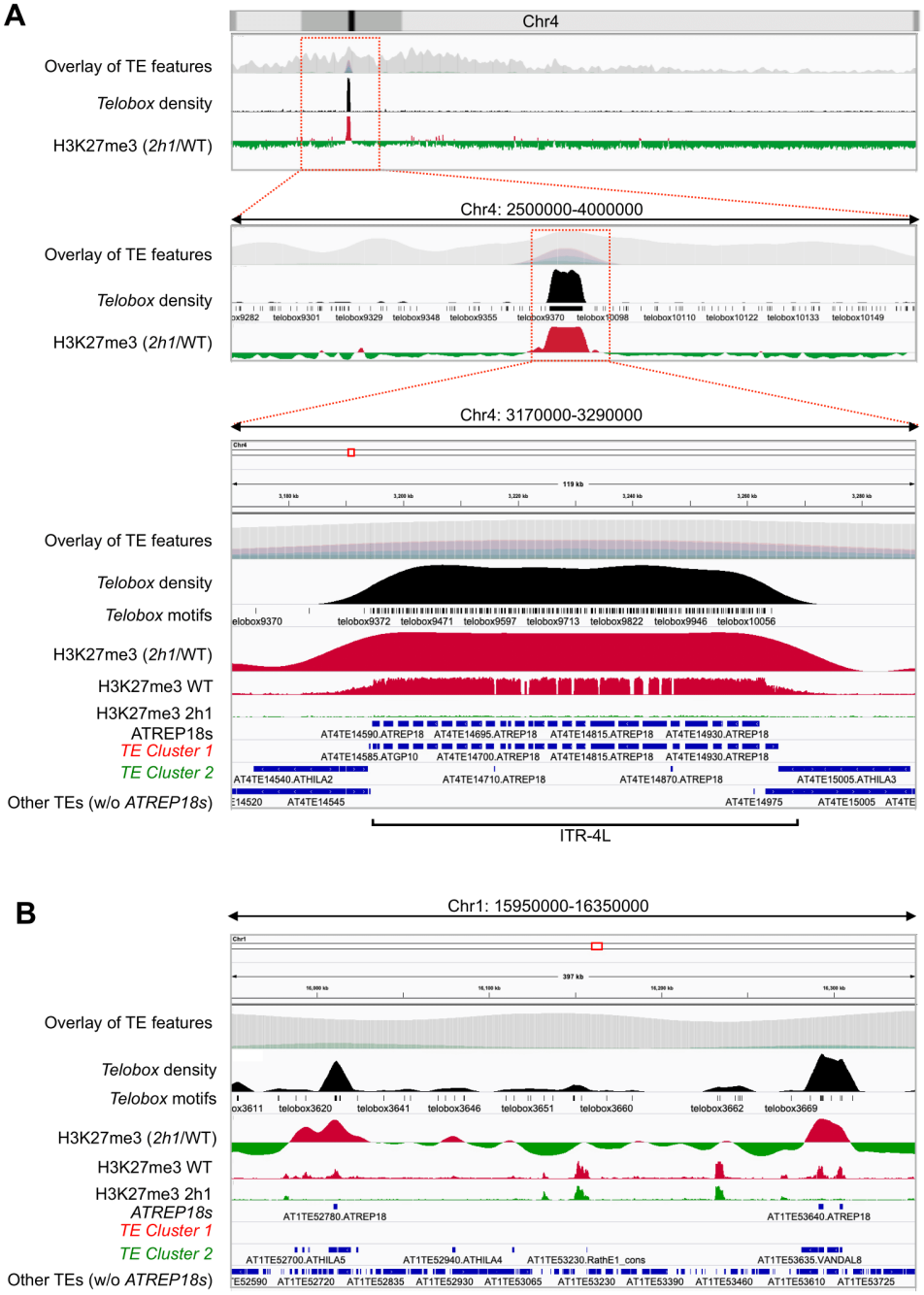
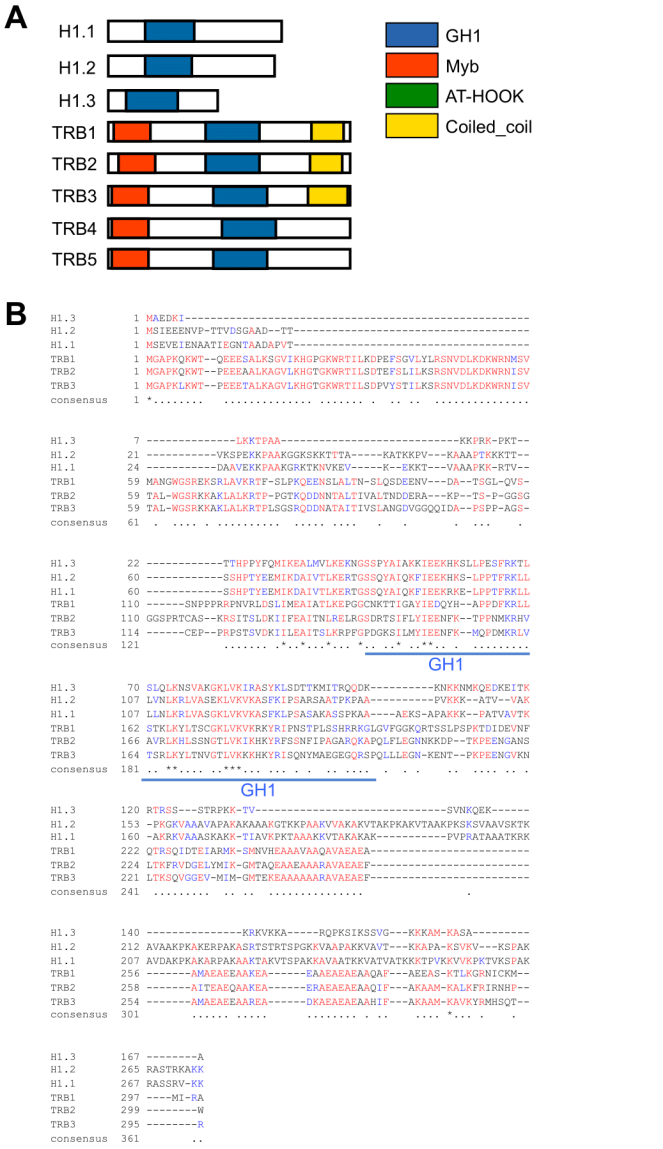



**Figure S11. Visualization of H3K27me3 profiles over *TE clusters 1-2* and *ATREP18* elements. A.** Chromosome 4 distribution of the indicated TEs and *ATREP18* elements, *telobox* motif distribution, and H3K27me3 levels. Top, middle and bottom panels are as in Figure 4A. **B.** Close-up view on some interspersed H3K27me3-enriched repeats of *TE cluster 2*, exemplifying a physical correlation between H3K27me3 enrichment and *telobox*-rich domains in *2h1* pericentromeres.

Figure S12




| 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 identity between TRB1, TRB2 and TRB3 proteins and the three H1 variants.

Figure S13

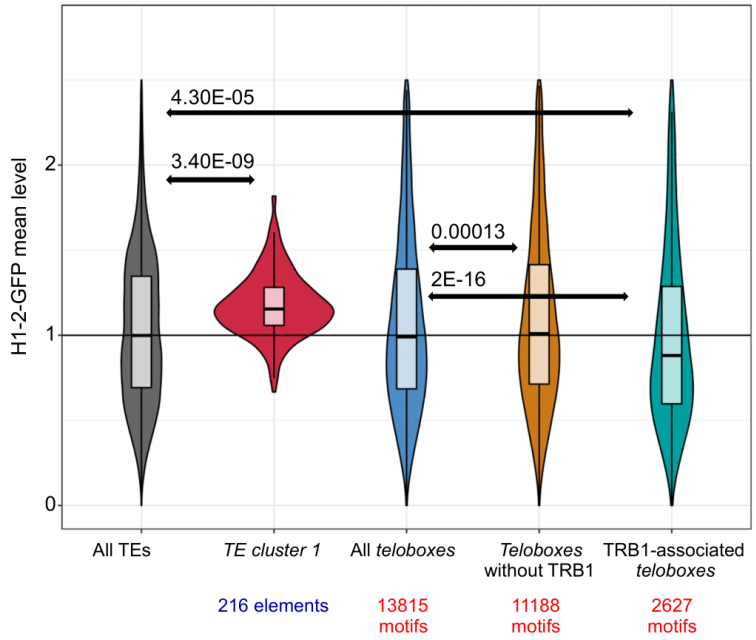



**Figure S13. *TE Cluster 1* displays elevated H1 level as compared to the ensemble of *Arabidopsis* TEs and to other *telobox*-containing regions.** In agreement with the proposed mutually exclusive binding of H1 and TRB1 over *teloboxes*, the set of *teloboxes* matching a known TRB1 peak from ChIP-seq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

Figure S14

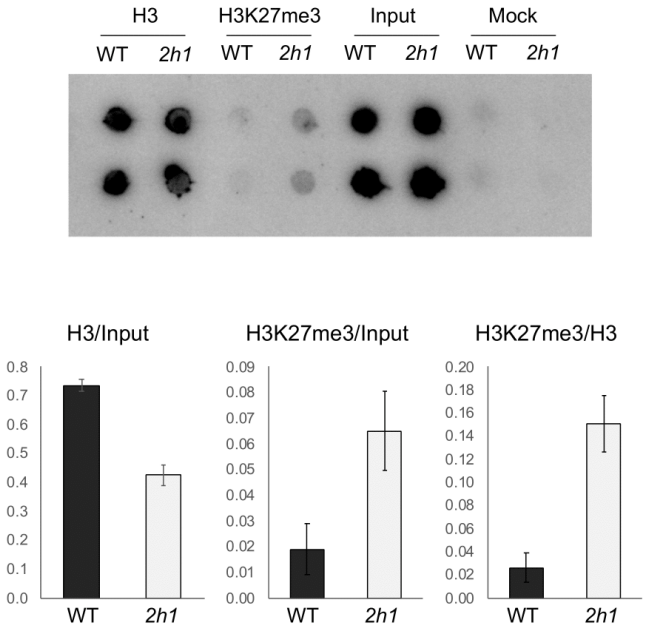



**Figure S14. Independent biological replicate of H3K27me3 and H3 ChIP DNA hybridization to telomeric probes completing Figure 6A.**

Figure S15

**A**

| | Genotype | Sample type | Replicate name | Enzyme | Total sequenced reads | Dangling end | Religation pairs | Self circle | % PCR duplicate | Valid interactions |
|---|---|---|---|---|---|---|---|---|---|---|
| This Study | WT Col-0 | Cotyledons | WT_rep1 | DpnII | 461075238 | 93890776 | 50150258 | 296024 | 25 | 227605683 |
| | | | WT_rep2 | DpnII | 488125661 | 107576431 | 47173304 | 286500 | 20 | 201684066 |
| | | | WT_rep3 | DpnII | 440385819 | 65273261 | 26305213 | 449415 | 32 | 300201639 |
| | *2h1* | Cotyledons | 2h1_rep1 | DpnII | 452087781 | 121437472 | 48027700 | 257520 | 19 | 158968785 |
| | | | 2h1_rep2 | DpnII | 448494540 | 57120253 | 26068664 | 424083 | 31 | 304266938 |
| | | | 2h1_rep3 | DpnII | 445283552 | 85226884 | 33019312 | 500832 | 26 | 240289837 |
| Liu et al 2016 | WT Col-0 | Entire seedlings | Liu_rep1 | DpnII | 143406012 | 35827356 | 8982912 | 1110590 | 17 | 69657043 |
| | | | Liu_rep2 | DpnII | 160112293 | 31038473 | 8095915 | 873572 | 19 | 80536362 |
| Wang et al 2016 | WT Col-0 | Entire seedlings | Wang_rep1 | DpnII | 247618590 | 32238398 | 9134591 | 1914560 | 29 | 146107616 |
| | | | Wang_rep2 | DpnII | 239477694 | 42804562 | 12643414 | 726839 | 25 | 141113166 |

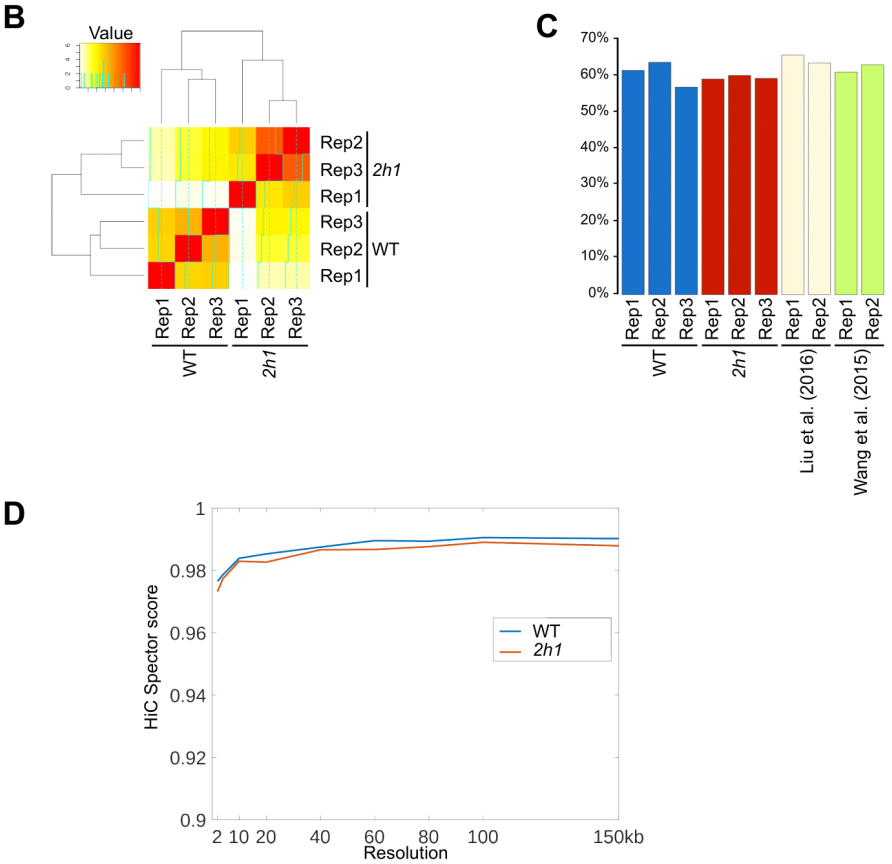



**Figure S15: Hi-C mapping and similarity index quality. A**. Mapping results of Hi-C libraries and comparison with re-processed Hi-C datasets from Liu et al. (2016) and Wang et al. (2015). **B**. Heatmap of similarity scores among WT and *2h1* mutant datasets calculated using spearman rank correlation between each sample in 100 kb bins. **C**. Comparison of intra-chromosomal reads over the total number of valid interactions among our samples and published DpnII-based Hi-C datasets from Liu et al. (2016) and Wang et al. (2015). Proportion of *cis* (intra-chromosomal) interactions among valid interactions is positively correlated with library quality for *in situ* Hi-C (Sun et al., 2020). **D.** Estimation of the Hi-C resolution achieved in this study. The curves show the HiC Spector score of chromosome 1 contact map against itself 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, using an Eigen value of 10, against the original Hi-C map with a resolution from 2kb to 150kb. For each resolution, the maximum HiC Spector score of the 30 down-samplings is reported. This analysis was performed using the merge of three independent biological replicates.

Figure S16

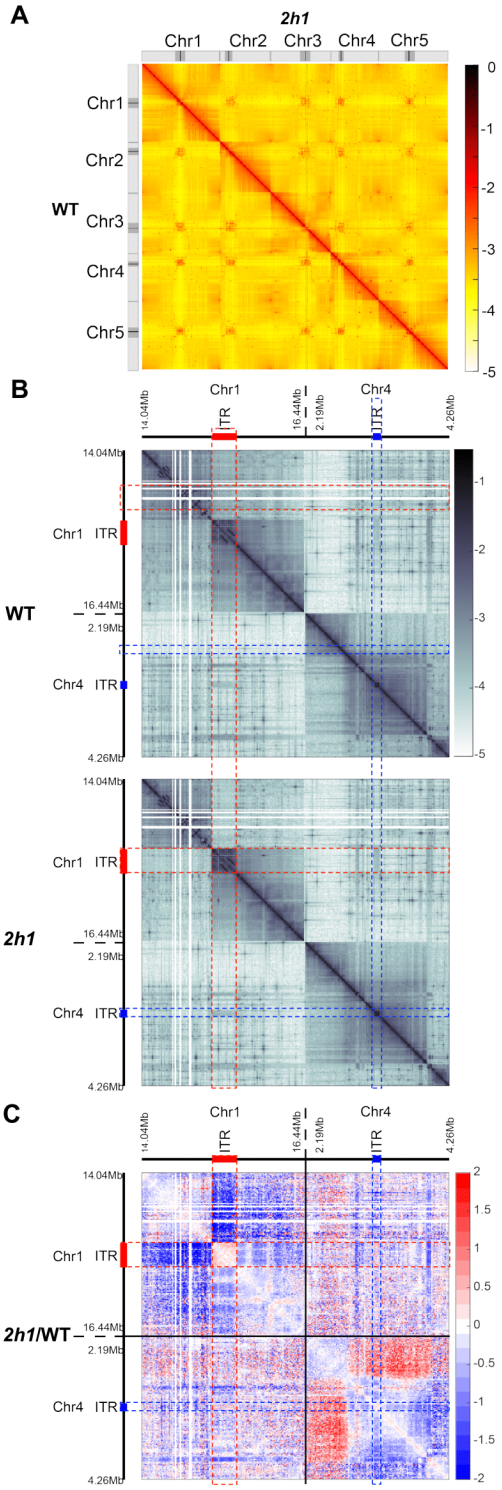



**Figure S16. Comparison of chromatin topology of WT and *2h1* nuclei. A.** Hi-C interaction frequency heatmap showing normalized $Log_{10}$ contact count at 100 kb resolution for all *Arabidopsis* chromosomes in WT (below the diagonal) and *2h1* (above the diagonal) nuclei. Lateral tracks show the positions of centromeres (black) and pericentromeres (gray). **B.** Same analysis than (A) illustrating interaction frequencies between the chromosomal regions spanning 1 Mb around ITR-1R (Chr1:15086191-15441067) and ITR-4L (Chr4:3192760-3265098) coordinates at a 10 kb resolution. **C.** Relative differences of interaction frequency between WT and *2h1* nuclei for the ITR regions illustrated in (B). $Log_2$ ratios of normalized interaction frequencies of *2h1* vs WT nuclei are shown at a 10kb resolution. All Hi-C data were analyzed using the merge of three independent biological replicates.

Figure S17

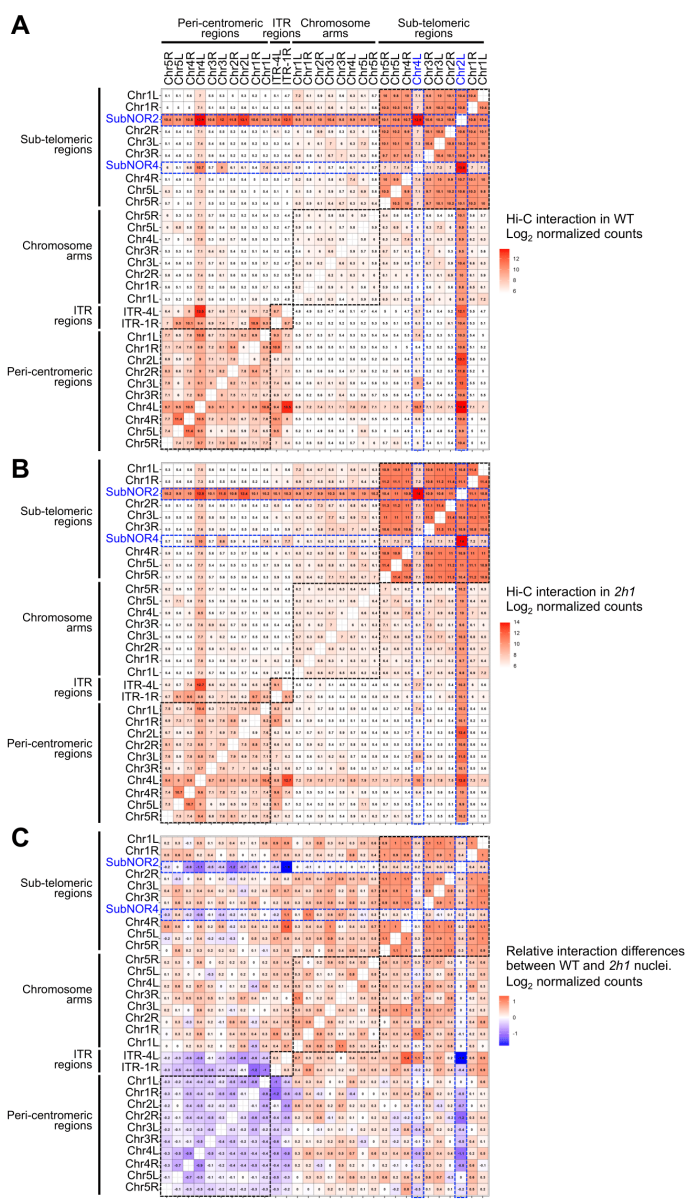
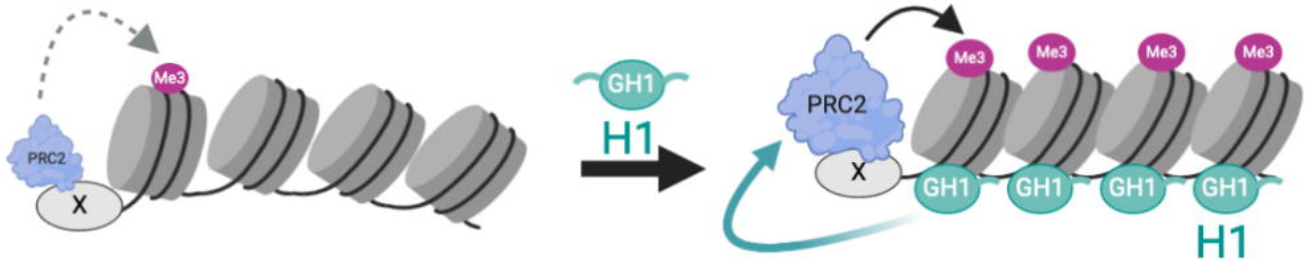



**Figure S17. Frequency of interaction between pericentromeric, ITR-1R/4L, and telomeric regions in WT and *2h1* nuclei, and difference between the two genotypes.** We probed four groups of regions: 1) ITR-1R and 4L coordinates, 2) sub-telomeric regions defined as the 100 kb terminal chromosomal regions adjacent to the telomeres, 3) pericentromeric regions represented by 100-kb segments located at 1 Mb from centromeres, and 4) 100-kb chromosome arm regions located at 5 Mb from the telomere positions. The sub-telomeric regions of chromosome 2 and 4 left arms are separated from the telomeres by the *NOR2* and *NOR4*, and therefore referred to as subNOR2 and SubNOR4 respectively (blue label). **A-B.** As highlighted by dashed lines, regions belonging to the same group tend to interact frequently together in both WT and *2h1* mutant plants. Accordingly, pericentromere-embedded ITR-1R and 4L frequently associate with other pericentromeric regions through inter-chromosomal contacts, but less frequently with all other probed groups except the *NOR2* region. Besides frequent long-range interactions with *NOR4*, the *NOR2*-adjacent Chr2L region also shows frequent interactions with diverse genome regions including ITR-4L and its neighboring pericentromeric regions on Chr4L. **C.** Relative differences of interaction frequency. With the exception of *NOR*-associated Chr2L and Chr4L regions, telomere-telomere and telomere-ITR interactions tend to increase in the absence of H1. Similarly, interactions between ITR-1R and 4L are slightly more frequent in the mutant line (see also Figure S16C). In contrast, interactions between ITRs and all pericentromeres tend to be reduced in the mutant line. The decrease in frequency of interaction is particularly marked between the *NOR*-adjacent regions and the pericentromeres of chromosome 2 and 4. 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. The indicated values are the median of three biological replicates (A and B) and the ratio of the medians (C).

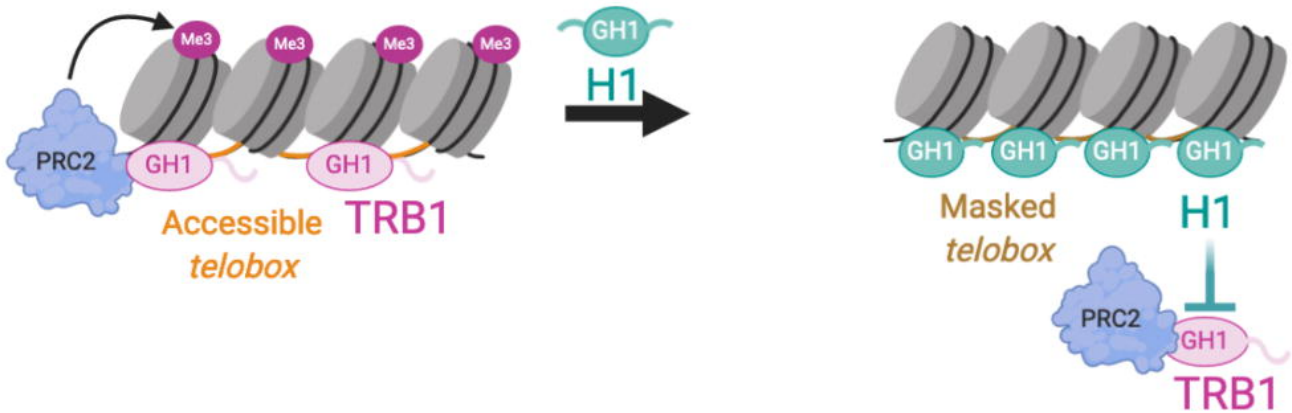
**H1 promotes H3K27me3 enrichment and limits chromatin accessibility of protein coding-genes**

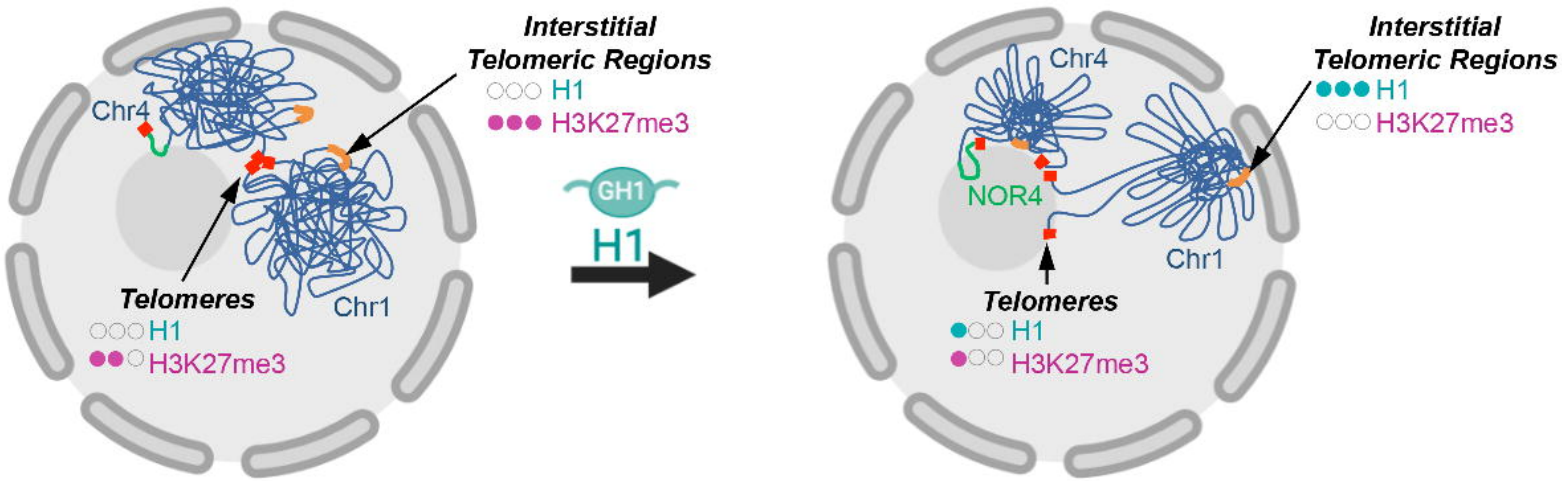
PRC2 target genes




**H1 prevents the invasion of clustered telomeric repeats by TRB1 and H3K27me3**

Telomeric repeats




**H1 triggers telomere individualization at the nucleolar periphery and orchestrates 3D genome organization**

Chromosome organization