MISS MINERVA SUSANA TREJO-ARELLANO (Orcid ID : 0000-0002-1982-3475)

DR LARS HENNIG (Orcid ID : 0000-0002-6645-1862)

Article type : Original Article

Arabidopsis Chromatin Assembly Factor 1 is required for occupancy and position of a subset of nucleosomes

Rafael Muñoz-Viana[1], Thomas Wildhaber[2], Minerva S. Trejo-Arellano[1], Iva Mozgová[1,3] and Lars Hennig[1]

Affiliations

[1] Swedish University of Agricultural Sciences, Department of Plant Biology and Linnean Center for Plant Biology, PO-Box 7080, SE-75007 Uppsala, Sweden.

[2] ETH Zürich, Department of Biology, Universitätsstrasse 2, CH-8092 Zürich, Switzerland.

[3] Institute of Microbiology, Centre Algatech, Opatovický mlýn, 37981 Třeboň, Czech Republic

Corresponding author: Lars Hennig (lars.hennig@slu.se)

Running title: Effects of loss of CAF-1 on nucleosomes in Arabidopsis

Keywords: Arabidopsis thaliana, chromatin, CAF-1, plant defense, GSE87421

Summary

CAF-1 is a major nucleosome assembly complex, which functions particularly during DNA replication and repair. Here, we studied how the nucleosome landscape changes in a CAF-1 mutant in the model plant *Arabidopsis thaliana*. Globally, most nucleosomes were not affected by loss of CAF-1 indicating the presence of efficient alternative nucleosome assemblers. Nucleosomes that we found depleted in the CAF-1 mutant were enriched in non-transcribed regions, consistent with the notion that CAF-1-independent nucleosome assembly can compensate for loss of CAF-1 mainly in transcribed regions. Depleted nucleosomes were particularly enriched in proximal promoters suggesting that CAF-1-independent nucleosome assembly mechanisms are often not efficient upstream of TSSs. Genes related to plant defense were particularly prone to loose nucleosomes in their promoters upon CAF-1 depletion. Reduced nucleosome occupancy at promoters of many defense-related genes is associated with a primed gene expression state that may considerably increase plant fitness by facilitating plant defense. Together, our results establish that the nucleosome landscape in Arabidopsis is surprisingly robust even in the absence of the dedicated nucleosome assembly machinery CAF-1 and that CAF-1 independent nucleosome assembly mechanisms are less efficient in particular genome regions.

Introduction

Nuclear DNA is packaged with histone proteins into nucleosomes forming the basic unit of chromatin. Nucleosome formation is catalyzed by replication-dependent and -independent mechanisms. CHROMATIN ASSEMBLY FACTOR 1 (CAF-1) is the major chromatin assembler for H3-H4 during replication (Stillman 1986, Smith et al.,

1989). Direct interaction with PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA) targets CAF-1 to sites of replication (Shibahara et al., 1999). CAF-1 is conserved in all eukaryotes and consists of three subunits called p150, p60 and p48 in mammals (Kaufman et al., 1995) and FASCIATA 1 (FAS1), FASCIATA 2 (FAS2) and MULTICOPY SUPPRESOR OF IRA (MSI1) in plants (Kaya et al., 2001). The two larger subunits FAS1/p150 and FAS2/p60 function exclusively in CAF-1 while the smallest subunit MSI1/p48 is a histone chaperone that functions in diverse complexes involved in chromatin formation, modification or remodeling (Hennig et al., 2005). CAF-1 binds to two H3-H4 dimers and promotes formation of $(H3\text{-}H4)_2$ tetramers (Winkler et al., 2012). Animals and plants possess specific histone H3 variants for replication-dependent and -independent chromatin assembly (Talbert et al., 2010); and human CAF-1 binds *in vivo* only to the replication-specific histone variant H3.1 but not to the replication-independent histone variant H3.3 (Tagami et al., 2004). In contrast, the HIRA histone chaperone normally binds mostly to H3.3 but can bind H3.1 in cells depleted of CAF-1 (Tagami et al., 2004, Ray-Gallet et al., 2011). Also in plants and yeast, HIRA partially compensates for loss of CAF-1 (Kaufman et al., 1998, Duc et al., 2015).

CAF-1 is essential in animals and causes delays in cell cycle and increased UV-sensitivity in yeast (Kaufman et al., 1997, Hoek et al., 2003, Nabatiyan et al., 2004, Houlard et al., 2006, Song et al., 2007). CAF-1 function is required for normal development of multicellular organisms (Yu et al., 2015). In the model plant *Arabidopsis thaliana* (Arabidopsis), the absence of a functional CAF-1 complex is not lethal, as *fas1* and *fas2* mutants are viable (Reinholz 1966, Leyser et al., 1992) (for review see (Ramirez-Parra et al., 2007b). Mutants for the smallest subunit MSI1 are embryo-lethal but this is CAF-1-independent and results from the function of MSI1 in

Polycomb repressive complex 2 (Köhler et al., 2003, Guitton et al., 2004). Mutants for CAF-1 in Arabidopsis show several phenotypic defects, including stem fasciation, abnormal leaf and flower morphology, and disorganization of the shoot and root apical meristems but are viable (Reinholz 1966, Leyser et al., 1992, Kaya et al., 2001). Also in rice, CAF-1 is required for normal SAM development (Abe et al., 2008). In contrast to the increased SAM size in Arabidopsis CAF-1 mutants, rice CAF-1 mutants have a reduced SAM indicating that mechanisms of SAM maintenance differ between rice and Arabidopsis. Additionally, Arabidopsis CAF-1 affects heterochromatin, transcriptional gene silencing, endoreduplication, cell differentiation, cell cycle duration, homologous recombination and trichome development (Kaya et al., 2001, Endo et al., 2006, Exner et al., 2006, Kirik et al., 2006, Ono et al., 2006, Schönrock et al., 2006, Ramirez-Parra et al., 2007a, Abe et al., 2008, Chen et al., 2008, Exner et al., 2008). Loss of Arabidopsis CAF-1 causes also telomere shortening and reduction of ribosomal DNA clusters (Mozgova et al., 2010, Muchova et al., 2015, Havlova et al., 2016, Pavlistova et al., 2016).

The viability of plant CAF-1 mutants clearly demonstrates that compensating nucleosome assembly mechanisms exist, likely involving HIRA (Duc et al., 2015). This is consistent with results from mammalian cells where HIRA-dependent nucleosome gap filling was observed (Ray-Gallet et al., 2011). It has however been so far unknown, how the genome-wide nucleosome landscape differs between wild-type and CAF-1 deficient cells. Important open questions are: Is there a genome-wide loss of nucleosomes in CAF-1 mutants? Are there local regions with reduced nucleosome occupancy? Are changes a consequence of altered transcription? Here we describe MNase-seq experiments that establish specific nucleosome landscapes

in a CAF-1 mutant. We find that global nucleosome patters are mostly unchanged but that specific nucleosomes are depleted or shifted. Depleted nucleosomes are often in promoters, particularly in proximal promoters of defense-related genes.

Results and Discussion

Genome sequence loss is not widespread in the fas2 *CAF-1 mutant*

MNase digestion followed by sequencing (MNase-seq) is commonly used to obtain high-resolution genome-wide maps of nucleosomes (Henikoff et al., 2011). Here, we applied MNase-seq to compare the nucleosome landscape in the Arabidopsis CAF-1 mutant *fas2-4* to that in wild type. Arabidopsis CAF-1 mutants lose certain sequences from their genome such as 45S rDNA sequences (Mozgova et al., 2010), which could interfere with the interpretation of MNase-seq data. Because it was not known whether also other genome regions undergo selective loss in *fas2*, we first used genome-resequencing to probe for loss of genome regions in *fas2* plants of generation 6. First, we compared sequence coverage between wild type and *fas2* for 45S rDNA sequences, which are known to be lost in *fas2*, and for other tandem or dispersed repetitive sequences, which are stable in *fas2* (Mozgova et al., 2010). Consistent with the earlier report about 90% of the 45S rDNA reads were lost in *fas2* (Supporting Table S1). In contrast, the tandem 5S rDNA repeats, the 180-bp centromeric repeats (Murata et al., 1994), dispersed repeats such as the retrotransposon *ATHILA* or the transposon *CACTA1*, as well as the single-copy controls *PP2A* and *ACTIN* had very similar coverage in wild type and the mutant arguing against loss of these sequences in *fas2*. Next, we divided the annotated genome into 1 kb bins and tested which regions have reduced coverage in *fas2*. Four such genome regions were identified comprising together 27 kb (Table 1). Two regions were long (10 kb and 15 kb) while the other regions were individual 1 kb

fragments. The 10 kb region with reduced coverage in *fas2* includes TE-rich 45S rDNA-containing sequences bordering Nucleolus Organizer Region 2 (NOR2). Interestingly, the region bordering NOR4, which does not contain any 45S rDNA annotated units, was not detected as changed. The 15 kb region with reduced coverage in *fas2* is a pericentromeric TE-rich region on chromosome 3 that also contains annotated 5.8S and 18S rRNA genes, which was not known to be affected in *fas2*. The other regions with reduced coverage in *fas2* were located in genes. We note that the reference genome-based resequencing analysis can be sensitive to errors in the TAIR10 genome assembly. Nevertheless, the experiment established that genome sequence loss in *fas2* is not widespread but restricted to sequences located in a few selected regions, in particular 45S rDNA and a pericentromeric TE-rich, rDNA-containing region on chromosome 3. Independent experiments using quantitative PCR on a subset of genomic loci confirmed the results of the sequencing for all but one of the individual 1 kb fragment regions (Figure 1). All regions with reduced coverage in *fas2* were excluded from the analysis of nucleosome occupation.

Nucleosomes with reduced occupancy in the fas2 *CAF-1 mutant are enriched in non-transcribed genome regions*

In contrast to the assumption that all nucleosomes are similarly digested by MNase, nucleosomes containing particular histone variants are much more sensitive to this treatment (Xi et al., 2011). H3.3- or H2A.Z-containing nucleosomes, for instance, are more fragile (Jin et al., 2007, Jin et al., 2009). It has been shown that chemical cross-linking greatly suppresses the high sensitivity of most unstable nucleosomes to MNase (Xi et al., 2011). Here, we used MNase digestion of cross-linked chromatin to

reduce loss of fragile nucleosomes. DNA from duplicates of MNase digested cross-linked chromatin of Col wild type and *fas2* was subjected to Illumina sequencing. For chromatin extraction, leaf number 6 was harvested from 38-days-old plants. At this developmental stage, leaf number 6 is fully expanded and all cell division and endoreduplication has ceased (Baerenfaller et al., 2012). Observed changes in nucleosomes are therefore not transient defects present only during or shortly after replication but are more permanent chromatin alterations that cannot be easily compensated for by CAF-1-independent chromatin assembly pathways. Presence of nucleosomes and changes in nucleosome position or occupancy were analyzed using DANPOS (Dynamic Analysis of Nucleosome Position and Occupancy by Sequencing) software (Chen et al., 2013) (see Experimental procedures). After excluding regions of low genomic coverage (see above), DANPOS identified 563,351 nucleosomes in the reference wild-type Col genome, which corresponds to a genome-wide average density of 1 identified nucleosome per 210 annotated base pairs. As an earlier study reported similar nucleosome spacing of 180-190 bp (Zhang et al., 2015), we conclude that most nucleosomes in Arabidopsis were identified in our approach.

Because CAF-1 assembles nucleosomes, we first focused on depleted nucleosomes, i.e. nucleosomes that had reduced coverage in *fas2*, which in extreme cases can be below detection limit. There were 128,197 such depleted nucleosomes, about one quarter of all detected nucleosomes. This demonstrates that changes in nucleosome occupancy are common in *fas2* but do not affect the entire genome. Independent experiments using H3 chromatin immunoprecipitation to evaluate nucleosome occupancy (see Experimental Procedures for details)

confirmed the results of the Mnase-seq: While six out of eight regions with depleted nucleosomes had also reduced occupancy in independent experiments, nucleosome occupancy was unchanged for all four control regions (Supporting Figure 1). Next, we asked whether depleted nucleosomes are enriched in certain genome regions. Because it is possible that nucleosome assembly pathways that compensate for loss of CAF-1 are related to transcription, we analyzed protein-coding gene bodies (i.e. the transcriptional units from transcriptional start site TSS to the transcriptional termination site TTS), proximal promoters (-200-TSS), transposable element (TE) gene bodies and intergenic sequences. Genes that are active or inactive in corresponding leaf tissue as reported in (Mozgova et al., 2015) (see Experimental procedures for details) were analyzed separately (Table 2). Depleted nucleosomes were detected in all analyzed genome features, but they were significantly more often found in non-transcribed regions such as proximal promoters and intergenic sequences than expected by chance (5.4% instead of 2.5% and 44.4% instead of 35.1%, respectively). Conversely, depleted nucleosomes were less often found on gene bodies than expected by chance (36.5% instead of 45.4% and 7.4% instead of 9.8% for active and inactive genes, respectively). Interestingly, active and inactive protein-coding genes as well as mostly silent TE-genes all showed a similar reduction of about 0.8 in depleted nucleosomes, which strongly contrasted with the increase in depleted nucleosomes at promoters and in intergenic regions (Table 2). This result suggests that even occasional transcription can lead to reestablished nucleosome occupancy. Thus, lack of CAF-1 leads to reduced nucleosome occupancy preferentially in non-transcribed genome regions. Besides the nucleosomes with reduced occupancy, about 8000 nucleosomes had increased occupancy in *fas2*. These nucleosomes were significantly enriched in bodies of

active and inactive genes as well as in TE genes (Supporting Table S2). Note that the enrichment was most significant for TE genes (p=2.87E-173), followed by inactive genes (p=1.26E-28) and least significant for active genes (p=3.47E-05) suggesting that increased nucleosome occupancy in *fas2* is enriched at transcription units but is not a direct consequence of ongoing transcription. In contrast, nucleosomes with increased occupancy in *fas2* were underrepresented in non-transcribed genome regions such as promoters and intergenic sequences.

Because some repetitive genome regions are selectively lost in Arabidopsis CAF-1 mutants, we investigated nucleosome occupancy on telomeric and centromeric repeats, on 5S rDNA and on 45S rDNA blocks (Supporting Figure S2) using a strategy proposed in (Schwartz et al., 2016). Read coverage indicated presence of positioned nucleosomes in wild type on telomeric and centromeric repeats but not on 5S rDNA or 45S rDNA blocks. In *fas2*, telomeres generated much less signal that did not indicate positioned nucleosomes suggesting that telomeric chromatin differs substantially between wild type and *fas2*. Similarly, centromeric repeats generated much less signal in *fas2* but the nucleosomal pattern was not lost. The analyzed 5S rDNA block contained a nucleosome free region at the 5´end, which was lost in *fas2* indicating that nucleosomal coverage spread across the locus. No changes in nucleosome occupancy were obvious at the 45S rDNA. Together, nucleosome occupancy was changed on some repetitive regions in *fas2* mutant plants.

Next, we asked whether there are any particular features characteristic for nucleosomes with reduced coverage such as the presence of DNA and histone

methylation, or the enrichment in canonical and variant histones. Chromatin scores were calculated for DNA surrounding identified nucleosome dyads in wild type using 19 genome-wide epigenome datasets (see Experimental Procedures for details). Although medians for all scores differed significantly between nucleosomes with reduced occupancy and other nucleosomes (Wilcoxon signed-rank test, $p<0.05$), which is consistent with the non-random distribution of affected nucleosome in the genome, the amplitude of the difference was always very small (Supporting Figure S3). Among the small differences, the largest effects were observed for H1 and H3K4me2, which were less abundant at wild-type nucleosomes that have reduced occupancy in *fas2*, as well as for H3.3 and H3K27me3, which were more abundant at such nucleosomes. It was unexpected to find increased H3.3 signals; however, this is consistent with the enrichment of not only nucleosomes with reduced occupancy (Table 2) but also H3.3 and H3K27me3 at many promoters (Shu et al., 2014). The observation that H3.3 nucleosomes were not specifically protected from being lost in *fas2* may indicate that efficient histone replacement with H3.3 depends on undisturbed structure and composition of the chromatin substrate. Other, smaller differences were observed for acetylated H3 (H3K9ac, H3K27ac, H3K36ac), again consistent with the presence at promoters or around the TSS. Signals for H2A.Z, H3K4me1, H3K4me3, H3K36me2, H3K36me3, H3S10p, H3S28p and DNA methylation differed least in this comparison. Together, these results argue against a specific strong epigenetic signal steering changes in nucleosome occupancy in *fas2*.

Reduced nucleosome occupancy in *fas2* preferentially in non-transcribed genome regions is consistent with the notion that CAF-1-independent nucleosome assembly is most efficient in transcribed genome regions (Gurard-Levin et al., 2014).

However, even in transcribed genome regions many nucleosomes were significantly depleted, suggesting that even in these regions CAF-1-independent nucleosome assembly can only partially restore a normal nucleosome landscape. Earlier work had shown that transcriptional changes in this CAF-1 mutant tissue are rather limited (Mozgova et al., 2015), indicating that the altered nucleosome landscape is not a direct and trivial consequence of altered gene expression. In addition, we observed maintenance of nucleosomal occupancy even in non-transcribed regions, strongly implying that CAF-1-independent nucleosome assembly mechanisms are not strictly dependent on transcription. This notion is consistent with the observation that the frequency of depleted nucleosomes did not differ considerably between active and inactive genes (18.3% and 17.1% of the detectable nucleosomes, respectively). Alternatively, prevalent genome-wide background transcription may suffice to reestablish nucleosomes even outside of annotated transcription units by a transcription-coupled nucleosome assembly process.

Next, we analyzed the distribution of depleted nucleosomes around the TSS of all protein-coding genes. Nucleosomes were binned according to their distance to the nearest TSS and p-values for the presence of depleted nucleosomes were calculated per bin (Figure 2A). Depleted nucleosomes were observed more often than expected immediately upstream of the TSS, corresponding to the -1 nucleosome. Conversely, depleted nucleosomes were observed less often than expected downstream of the TSS. Notably, the same pattern of nucleosome depletion was observed for genes with unchanged (Figure 2B) or with changed expression in *fas2* (Figure 2C), suggesting that the observed change in nucleosome

occupancy is not a trivial consequence of transcriptional changes in *fas2*. We note that the number of genes mis-regulated in *fas2* is relatively small and that the majority of genes with nucleosome depletion are not changed in expression. It was possible that nucleosome depletion affects genes independent of gene function. Alternatively, specific functional groups might be more affected than others. Gene ontology (GO) enrichment analysis for genes with the most extreme nucleosomal changes suggested that genes were not affected randomly. Instead, genes with functions related to plant defense signaling had more often depleted nucleosomes in their promoters than other genes (Table 3). Of the 19 most significantly enriched categories ($p<0.05$), 13 were related to plant defense. Because nucleosome depletion was most significant immediately upstream of the TSS, we focused on genes with depleted nucleosomes in the proximal promoter (-200bp – TSS). Again, a similar enrichment of defense-related GO categories was observed (Table 3, Supporting Table S3). Together, nucleosome depletion in *fas2* was most pronounced in proximal promoters and preferentially affected specific functional groups of genes.

The frequent localization of depleted nucleosomes in proximal promoters suggested that CAF-1-independent nucleosome assembly mechanisms are often not efficient upstream of TSSs. The result that genes related to plant defense were particularly prone to loose nucleosomes in their promoters is consistent with earlier findings that CAF-1 deficiency causes hypersensitivity of some defense related genes in Arabidopsis (Mozgova et al., 2015). It was found that the CAF-1 mutant state resembles a naturally hypersensitive (primed) state of these genes (Mozgova et al., 2015). The genome-wide analysis performed here established that reduced nucleosome occupancy in promoters is not limited to the selected defense genes tested earlier but is a global effect for this class of genes. The results reveal also that

nucleosomes in promoters of genes belonging to other functional classes are not as often depleted. Thus it appears that CAF-1-independent nucleosome assembly mechanisms are less efficient in promoters of defense-related genes than in other promoters. It is possible that CAF-1-independent nucleosome assembly mechanisms are repelled from promoters of defense-related genes. Alternatively, CAF-1-independent nucleosome assembly mechanisms may function at promoters of most genes but low-level defense signaling may expel nucleosomes from promoters of defense-related genes.

Because defense priming is of great importance for plant fitness, the mechanism that restricts CAF-1-independent nucleosome assembly at promoters of many defense-related genes to allow establishment and maintenance of a primed gene expression state is of considerable physiological relevance.

Nucleosome shifts in the* fas2 *CAF-1 mutant are mainly in gene bodies

Nucleosomes may not only change occupancy but also shift along the DNA. We detected 25,864 significantly shifted nucleosomes in *fas2* (see Experimental Procedures and Supporting Figure S4 for details). The functional impact of nucleosome shift depends on whether shifted nucleosomes maintain rotational position, i.e. whether the alternative translational positions change accessibility of bases inside nucleosomal DNA. Prevalent maintenance of rotational positions would be reflected in the preference for shifts by 10-11 bp steps (Luger et al., 1997). However, when we analyzed the distribution of shift amplitudes no evidence for any such preference could be seen (Supporting Figure S5); the 7.9% of the shifts that maintain rotational positions reflect the fraction expected by chance. Thus,

nucleosome shifts in *fas2* have no obvious tendency to maintain or interrupt rotational position. In contrast to depleted nucleosomes, which were enriched in non-transcribed regions such as proximal promoters and intergenic sequences, shifted nucleosomes were enriched in gene bodies of protein coding as well as of TE genes (Table 4). Notably, both active and inactive as well as mostly silent TE genes showed enrichment of shifted nucleosomes. In contrast, proximal promoters and intergenic sequences had less shifted nucleosomes than expected by chance. Next, we wondered whether similar to nucleosome depletion in proximal promoters, nucleosome shifts in gene bodies affected specific functional groups of genes. Indeed, GO enrichment analysis revealed that seven of the ten enriched categories (70%) were related to development (Table 5). Nucleosomes on genes can be shifted in the sense (forward) or antisense (reverse) direction. Similar numbers of nucleosomes were shifted forward or reverse in *fas2* (12847 vs. 13017). Next, we analyzed the position of shifted nucleosomes on genes in more detail. Nucleosomes were binned according to their distance to the nearest TSS and p-values for reduction or enrichment of shifted nucleosomes were calculated per bin (Figure 3A). As noticed above, shifted nucleosomes were enriched downstream of the TSS and reduced upstream. This was independent of the shift direction. However, the forward shifted nucleosomes showed the strongest enrichment 100 bp downstream of the TSS corresponding to the +1 nucleosome while reverse shifted nucleosomes were reduced at this position and had the strongest enrichment at 400 bp downstream of the TSS. The same pattern of local enrichment of shifted nucleosomes was observed when only using genes with unchanged expression in *fas2* (Figure 3B), suggesting that similar to the observed changes in nucleosome occupancy the nucleosome shifts are not a trivial consequence of transcriptional changes in *fas2*.

Conclusions

This study revealed that nucleosomes that are depleted in an Arabidopsis CAF-1 mutant are enriched in non-transcribed regions. This result is consistent with the notion that CAF-1-independent nucleosome assembly is most efficient in transcribed genome regions. Because the analyzed tissue did not contain replicating nuclei, additional chromatin changes may exist in CAF-1 mutants transiently during or after replication but were masked by post-replicative CAF-1-independent nucleosome assembly. Efficiency of CAF-1-independent nucleosome assembly at promoters differs for different functional groups of genes indicating the action of a not well understood mechanism of gene regulation. The mechanism limiting CAF-1-independent nucleosome assembly at promoters of many defense-related genes is important for plant fitness as it allows establishment and maintenance of a primed gene expression state that contributes to an efficient defense against pathogens.

Experimental procedures

Plant material and growth conditions

The *Arabidopsis thaliana* wild-type and T-DNA insertion lines were in the Columbia-0 (Col) background; *fas2-4* (SALK_033228) was described before (Exner et al., 2006). Because CAF-1 mutants suffer from a progressive loss of tandem repeats and vary between generations (Mozgova et al., 2010), all adult-plant experiments were performed on *fas2*-4 generation 6 (G6) plants. Seeds were sown on 0.5× basal salts Murashige and Skoog (MS) medium (Duchefa, Haarlem, The Netherlands), stratified at 4ºC for 48 h, and allowed to germinate for 10 days in growth chambers under short day (8 h light, 100 µmoles m^{-2} s^{-1}, 22 °C) conditions. Plantlets were planted in

soil and grown in growth chambers under the same conditions. For chromatin extraction, leaf number 6 was harvested after 38 d at ZT (zeitgeber time) = 7 h. At this time, leaf number 6 is fully expanded and all cell division and endoreduplication has ceased (Baerenfaller et al., 2012). Note that this condition corresponds to standard conditions defined in the AGRON-OMICS project that have been used in several studies facilitating data comparison (Baerenfaller et al., 2012, Derkacheva et al., 2013, Shu et al., 2013, Shu et al., 2014, Baerenfaller et al., 2015, Mozgova et al., 2015).

Genome resequencing

For genome resequencing, generation 6 (G6) *fas2*-4 and corresponding segregated wild-type control plants were used. 10-day old seedlings were harvested and DNA was isolated using a MagJET Plant Genomic DNA Kit (Thermo Fisher Scientific) according to the manufacturer´s instructions. Genomic DNA sequencing libraries were generated using a TruSeq DNA LT Library Preparation kit (Illumina Inc) and sequenced using 100 bp paired-end mode on an Illumina HiSeq 2000 system (Illumina Inc). Reads with total sequencing quality of less than 30 Phred or shorter than 10 nucleotides were discarded. For each library, 25 Mio clean paired-end reads were mapped to the reference genome TAIR10 using bwa with default parameters (Li et al., 2009b). The mapping efficiency was 94% and 87% for the Col and *fas2* mutant library, respectively. Duplicated paired mappings were removed using samtools rmdup (Li et al., 2009a). The annotated nuclear genome was divided into 1 Kb bins and the number of mapped reads falling into each bin were counted. Bins containing less reads in Col than the tenth lower quantile of all bins (67 reads per 1 Kb bin) were not considered for further analysis. A bin-wise one-sided binomial test was used to test for reduced coverage in *fas2*. Multiple testing correction was done

according to (Storey et al., 2003). Bins were considered to have reduced coverage in *fas2* if q<0.05 and log2FC < -1.5. To account for different numbers of mapped reads in wild type and *fas2*, coverage was expressed in FKPM (fragments per 1 kilobase and 1 million mapped fragments).

MNase-seq

Leaf material was cross-linked using 1% formaldehyde under vacuum for 15 min. Cross-linking was quenched using an excessive amount of glycine (0.125 M) and treatment with vacuum for another 5 min. The leaves were rinsed with water, blot-dried on filter paper and ground in liquid nitrogen. To extract chromatin, 100 mg of ground frozen leaf material was treated in Nuclei Extraction Buffer (NEB: 20 mM PIPES-KOH pH 7.6, 1 M hexylene glycol, 10 mM $MgCl_2$, 0.1 mM EGTA, 15 mM NaCl, 60 mM KCl, 0.5% Triton-X, 5 mM β-mercaptoethanol and EDTA-free protease inhibitor cocktail (Roche, Rotkreuz, Switzerland)) for 15 min at 4°C. The homogenate was filtered through CellTrics Nylon filter (50 µm) (Partec, Görlitz, Germany) and pellets were collected by centrifugation for 10 min at 1500x*g* at 4°C. Pellets were washed once in MNase buffer (50 mM Tris-HCl pH 8, 10 mM NaCl, 5 mM $CaCl_2$ and EDTA-free protease inhibitor cocktail (Roche)) and were re-suspended in 100 µl of MNase buffer. An aliquot of 20 µl was set aside as input control, and 20 µl of MNase buffer and 1.3 µl of RNase A (30 µg/µl, Sigma-Aldrich, Buchs, Switzerland) were added to the rest of the suspension. This mixture was digested with Micrococcal Nuclease (MNase, New England BioLabs, Schwalbach, Germany, final concentration 0.2 U/µl) for 8 min in MNase Buffer. The reaction was stopped with 10 mM EDTA. The supernatant after a centrifugation for 5 min at 1500x*g* at 4°C was

collected as supernatant 1. The pellet was re-suspended in S2 buffer (1 mM Tris-HCl pH 8, 0.2 mM EDTA and EDTA-free protease inhibitor cocktail (Roche)) and treated for 30 min on ice. The suspension was centrifuged again for 5 min at 1500x*g* at 4°C. After pooling supernatant 1 and 2, formaldehyde crosslinks were reversed by treatment at 65°C overnight with elevated salt concentration (1 M NaCl). The mixture was treated with de-proteinization solution (0.02 M EDTA; 0.1 M Tris-HCl pH 6.5; 1 µl Protease K (Sigma)) at 45°C for 3 hours. Nucleosome DNA was recovered as above. Digested DNA was resolved on 1% agarose gels. Gel slabs containing the DNA fraction 150 ± 50 bp were sliced and DNA was recovered using a GFX PCR DNA and Gel Band Purification Kit (Amersham Bioscience, Dübendorf, Switzerland) according to the manufacture´s protocol. Sequencing libraries were constructed using an Ovation Ultralow Library System (Nugen, San Carlos, CA), including 8 different barcodes as described in the manufacture´s protocol. Sequencing was performed on an Illumina HiSeq 2000 System in 100 bp paired-end mode using v3 sequencing chemistry at the ScienceForLife Laboratories in Uppsala, Sweden.

Bioinformatics analysis

FastQC v0.10.1 was used to check read quality. Libraries of 30.3 and 39.4 million paired reads for Col and *fas2*, respectively, were mapped with bowtie2 against the TAIR10 genome with efficiencies of 85% using the local option. Mapped MNase-seq DNA samples were analyzed with DANPOS (Chen et al., 2013), which extracted the position of nucleosome dyads and performed statistical tests to identify significantly changed nucleosomes. Briefly, DANPOS takes as input two paired-end mapped bam files, removes read duplicates, calculates the position of the nucleosome adjusting the forward and reverse reads to the middle of their positions and calculates

nucleosome occupancy as local read coverage. Then DANPOS compares the two samples by first normalizing samples using quantile normalization and second performing differential signal calculation with a Poisson test. DANPOS identifies the position of the nucleosome dyad. A nucleosome is considered as shifted when the position of the dyad in sample 1 is 50-90 nt away from the position in sample 2. Small differences in dyad position (<50 nt) are considered noise rather than a shift; large differences in dyad position (>90) are considered to reflect different nucleosomes rather than shifts. Supporting Figure S4 illustrates the method of assigning nucleosome shifts. For more details on algorithms and tests used by DANPOS see (Chen et al., 2013). Further analysis was performed with custom scripts in R v3.1.2. Gene Ontology analysis was performed using the GOstats R package. The hypergeometric test with Bonferroni correction was used with a filter requiring three genes with a manually assigned annotation as minimum category population. Transcriptome data are from (Mozgova et al., 2015) based on Affymetrix tiling arrays with a signal of 5 as threshold to differentiate active from inactive genes (Rehrauer et al., 2010). This study identified 307 upregulated and 65 downregulated genes in fully grown leaves of *fas2-4*, which corresponded to the material used here. Thus, the vast majority of genes is not differentially expressed in this material. For the epigenetic profiling of depleted and non-depleted nucleosomes, scores were calculated by averaging the Z-score coverage values +/-50 bp around the nucleosome dyad detected by DANPOS in WT conditions. Chromatin profiles were taken from the literature: H1.1 and H1.2 (Rutowicz et al., 2015); H3K27ac, H3K27me3, H3K4me3, H3K9ac, H3K9me2 and H3S10p (Baerenfaller et al., 2016); H3.3 (Shu et al., 2014); H2A.Z (Coleman-Derr et al., 2012); H3K4me1 (Zhang et al., 2009); H3K4me2 and H3K36me2 (Luo et al., 2013); H3K36ac and H3K36me3

(Mahrez et al., 2016); DNA inaccessibility (Shu et al., 2012) and DNA methylation (Stroud et al., 2013) (H1.1 and H1.2: 3-weeks-old plants; H3K27ac, H3K27me3, H3K4me3, H3K9ac, H3K9me2, H3S10p, H3K36ac, H3K36me3, H3.3 and DNA inaccessibility: Leaf number 6, 5 weeks old plants; H2A.Z: 4 weeks post germination roots; H3K4me and DNA methylation1: 3 weeks old plants; H3K4me2 and H3K36me2: 2 weeks old plants).

Experimental confirmation of loss of genome regions

Genomic DNA was isolated from 100 mg (fresh weight) 10-day-old seedlings using the MagJET Plant Genomic DNA Kit (Thermo Fisher Scientific) according to manufacturer´s instructions. Concentration of the purified RNA-free DNA was quantified using the Qubit dsDNA BR (broad range) Assay Kit (Thermo Fisher Scientific) and adjusted to 10 ng/µl in all samples. 0.02 – 20 ng of gDNA were used in 20 µl quantitative PCR reaction using 5x HOT FIREPol® EvaGreen® qPCR Supermix (Solis Biodyne) and primers specified in Supporting Table S4. Quantification was done using a MyiQTM Single Color Real Time PCR detection system (BIO-RAD), the results were normalized to *ACT2* (*At3G18780*) and to the wild-type seedlings. The experiment was performed in three biological replicates.

Chromatin immunoprecipitation

Col and *fas2* fully expanded leaves from 5-week-old plants grown under long-day conditions were collected and cross-linked using 1% formaldehyde for 10 minutes followed by quenching by 0.125 M glycine for 5 min under vacuum. 100 mg of ground material was used to extract crude nuclei as described above for MNase-sequencing. Nuclei were washed once using 200 µl of MNase buffer and collected

by centrifugation (1500×g, 10 min, 4°C). Extracted nuclei were re-suspended in 50 µl of MNase buffer and 10 µl of non-digested sample was taken as input control. 5 U of MNase (New England BioLabs) were added and the suspension was incubated for 6 minutes at 37°C. The reaction was stopped by adding EDTA (final 10 mM) and SDS (final 1%). 1 cycle of 30 sec sonication was applied (Bioruptor sonicator, Diagenode), the suspension was diluted to 500 µl using ChIP buffer (16.7 mM Tris-HCl - pH 8.0, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 1 x cOmplete EDTA-free proteinase inhibitor (Roche)), and the chromatin was cleared by centrifugation at 4500xg, 5 min, 4°C. Immunoprecipitation was performed as described (Mozgova et al., 2015). For each immunoprecipitiation, 200 µl of cleared chromatin were incubated with 1.5 µg of anti-histone H3 antibody (Upstate Millipore, cat. no. 07-690) or IgG (Sigma, cat. no. I5006) as control. Immunoprecipitated and input DNA was de-crosslinked, treated with Proteinase K (Sigma), extracted using phenol-chloroform and quantified in technical duplicates by quantitative PCR (qPCR) on the RotorGene RG-3000 system (Corbett Research) using 5X HOT FIREPol Eva Green qPCR Mix Plus (ROX) (Solis Biodyne) and primers specified in Supporting Table S5. DNA amount was related to input control before MNase digest. Genomic regions representing nucleosomes with reduced occupancy in *fas2* as identified by the MNase-seq approach were selected for confirmation by MNase-ChIP-qPCR. Only single-copy regions where the closest neighboring gene did not differ in expression level in *fas2* (Mozgova et al., 2015) were considered. Regions with no changes in nucleosome occupancy or position were taken as controls. Primers were designed to amplify regions spanning approx. 50 bp upstream and 50 bp downstream of the wild-type nucleosome dyad axis position. ChIP was performed in triplicates using leaves of independent plants; biological replicates were averaged and a lower anti-H3 DNA

recovery in *fas2* than in wild type was tested using a one-sided paired Student´s t-test.

Accession numbers

Data are available at GEO (accession number GSE87421).

Acknowledgment

This work was supported by grants from the Swiss National Science Foundation SNF, the Swedish Science Foundation VR and the Knut-and-Alice-Wallenberg Foundation. Authors confirm that they have no conflicts of interest to declare.

Short legends for Supporting Information

Supporting Table S1. Sequence coverage in selected repeat regions.

Supporting Table S2. Genome distribution of nucleosomes with increased occupancy in *fas2*.

Supporting Table S3. Enriched GO terms in genes with depleted nucleosomes in their proximal promoters (-200:TSS).

Supporting Table S4. Primers used for verification of genome resequencing using quantitative PCR.

Supporting Table S5. Primers used for verification of MNase-sequencing using quantitative PCR.

Supporting Figure S1. Validation of MNase-seq results.

Supporting Figure S2. Repetitive regions did not undergo selective loss of nucleosomes in *fas2*.

Supporting Figure S3. Epigenetic profile of depleted and non-depleted nucleosomes.

Supporting Figure S4. Scheme of the method for assigning nucleosome shifts.

Supporting Figure S5. The magnitude of the shift of displaced nucleosomes in *fas2* does not follow a periodic trend.

References:

Abe, M., Kuroshita, H., Umeda, M., Itoh, J. and Nagato, Y. (2008) The rice flattened shoot meristem, encoding CAF-1 p150 subunit, is required for meristem maintenance by regulating the cell-cycle period. *Dev Biol* **319**, 384-393.

Baerenfaller, K., Massonnet, C., Hennig, L., Russenberger, D., Sulpice, R., Walsh, S., Stitt, M., Granier, C. and Gruissem, W. (2015) A long photoperiod relaxes energy management in Arabidopsis leaf six. *Current Plant Biology* **2**, 34-45.

Baerenfaller, K., Massonnet, C., Walsh, S., Baginsky, S., Buhlmann, P., Hennig, L., Hirsch-Hoffmann, M., Howell, K.A., Kahlau, S., Radziejwoski, A., Russenberger, D., Rutishauser, D., Small, I., Stekhoven, D., Sulpice, R., Svozil, J., Wuyts, N. and Sti (2012) Systems-based analysis of Arabidopsis leaf growth reveals adaptation to water deficit. *Mol Syst Biol* **8**, 606.

Baerenfaller, K., Shu, H., Hirsch-Hoffmann, M., Futterer, J., Opitz, L., Rehrauer, H., Hennig, L. and Gruissem, W. (2016) Diurnal changes in the histone H3 signature H3K9ac|H3K27ac|H3S28p are associated with diurnal gene expression in Arabidopsis. *Plant Cell Environ* **39**, 2557-2569.

Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X. and Li, W. (2013) DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res* **23**, 341-351.

Chen, Z., Tan, J.L., Ingouff, M., Sundaresan, V. and Berger, F. (2008) Chromatin assembly factor 1 regulates the cell cycle but not cell fate during male gametogenesis in Arabidopsis thaliana. *Development* **135**, 65-73.

Coleman-Derr, D. and Zilberman, D. (2012) Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. *PLoS Genet* **8**, e1002988 LID-10.1371/journal.pgen.

Derkacheva, M., Steinbach, Y., Wildhaber, T., Mozgova, I., Mahrez, W., Nanni, P., Bischof, S., Gruissem, W. and Hennig, L. (2013) Arabidopsis MSI1 connects LHP1 to PRC2 complexes. *EMBO J* **32**, 2073-2085.

Duc, C., Benoit, M., Le Goff, S., Simon, L., Poulet, A., Cotterell, S., Tatout, C. and Probst, A.V. (2015) The histone chaperone complex HIR maintains nucleosome occupancy and counterbalances impaired histone deposition in CAF-1 complex mutants. *Plant J* **81**, 707-722.

Endo, M., Ishikawa, Y., Osakabe, K., Nakayama, S., Kaya, H., Araki, T., Shibahara, K.I., Abe, K., Ichikawa, H., Valentine, L., Hohn, B. and Toki, S. (2006) Increased frequency of homologous recombination and T-DNA integration in Arabidopsis CAF-1 mutants. *EMBO J* **25**, 5579-5590.

Exner, V., Gruissem, W. and Hennig, L. (2008) Control of trichome branching by Chromatin Assembly Factor-1. *BMC Plant Biol.* **8**, 54.

Exner, V., Taranto, P., Schönrock, N., Gruissem, W. and Hennig, L. (2006) Chromatin assembly factor CAF-1 is required for cellular differentiation during plant development. *Development* **133**, 4163-4172.

Guitton, A.E., Page, D.R., Chambrier, P., Lionnet, C., Faure, J.E., Grossniklaus, U. and Berger, F. (2004) Identification of new members of *FERTILISATION INDEPENDENT SEED* Polycomb group pathway involved in the control of seed development in *Arabidopsis thaliana*. *Development* **131**, 2971-2981.

Gurard-Levin, Z.A., Quivy, J.P. and Almouzni, G. (2014) Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annu Rev Biochem* **83**, 487-517.

Havlova, K., Dvorackova, M., Peiro, R., Abia, D., Mozgova, I., Vansacova, L., Gutierrez, C. and Fajkus, J. (2016) Variation of 45S rDNA intergenic spacers in Arabidopsis thaliana. *Plant Mol Biol*.

Henikoff, J.G., Belsky, J.A., Krassovsky, K., MacAlpine, D.M. and Henikoff, S. (2011) Epigenome characterization at single base-pair resolution. *Proc Natl Acad Sci U S A* **108**, 18318-18323.

Hennig, L., Bouveret, R. and Gruissem, W. (2005) MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes. *Trends Cell Biol* **15**, 295-302.

Hoek, M. and Stillman, B. (2003) Chromatin Assembly Factor-1 is essential and couples chromatin assembly to DNA replication in vivo. *Proc Natl Acad Sci U S A* **100**, 12183-12188.

Houlard, M., Berlivet, S., Probst, A.V., Quivy, J.P., Hery, P., Almouzni, G. and Gerard, M. (2006) CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. *PLoS Genet* **2**, 1686-1696.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K. and Felsenfeld, G. (2009) H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* **41**, 941-945.

Jin, C. and Felsenfeld, G. (2007) Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev* **21**, 1519-1529.

Köhler, C., Hennig, L., Bouveret, R., Gheyselinck, J., Grossniklaus, U. and Gruissem, W. (2003) Arabidopsis MSI1 is a component of the MEA/FIE Polycomb group complex and required for seed development. *EMBO J* **22**, 4804-4814.

Kaufman, P.D., Cohen, J.L. and Osley, M.A. (1998) Hir proteins are required for position-dependent gene silencing in Saccharomyces cerevisiae in the absence of chromatin assembly factor I. *Mol Cell Biol* **18**, 4793-4806.

Kaufman, P.D., Kobayashi, R., Kessler, N. and Stillman, B. (1995) The p150 and p60 subunits of chromatin assembly factor i: a molecular link between newly synthesized histones and DNA replication. *Cell* **81**, 1105-1114.

Kaufman, P.D., Kobayashi, R. and Stillman, B. (1997) Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking Chromatin Assembly Factor-1. *Genes Dev* **11**, 345-357.

Kaya, H., Shibahara, K., Taoka, K., Iwabuchi, M., Stillman, B. and Araki, T. (2001) *FASCIATA* genes for Chromatin Assembly Factor-1 in Arabidopsis maintain the cellular organization of apical meristems. *Cell* **104**, 131-142.

Kirik, A., Pecinka, A., Wendeler, E. and Reiss, B. (2006) The Chromatin Assembly Factor subunit FASCIATA1 is involved in homologous recombination in plants. *Plant Cell* **18**, 2431-2442.

Leyser, H.M. and Furner, I.J. (1992) Characterisation of three shoot apical meristem mutants of *Arabidopsis thaliana*. *Development* **116**, 397-403.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009a) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Li, H. and Durbin, R. (2009b) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260.

Luo, C., Sidote, D.J., Zhang, Y., Kerstetter, R.A., Michael, T.P. and Lam, E. (2013) Integrative analysis of chromatin states in Arabidopsis identified potential regulatory mechanisms for natural antisense transcript production. *Plant J* **73**, 77-90.

Mahrez, W., Trejo Arellano, M.S., Moreno-Romero, J., Nakamura, M., Shu, H., Nanni, P., Kohler, C., Gruissem, W. and Hennig, L. (2016) H3K36ac is an evolutionary conserved plant histone modification that marks active genes. *Plant Physiol*.

Mozgova, I., Mokros, P. and Fajkus, J. (2010) Dysfunction of Chromatin Assembly Factor-1 induces shortening of telomeres and loss of 45S rDNA in *Arabidopsis thaliana*. *Plant Cell* **22**, 2768-2780.

Mozgova, I., Wildhaber, T., Liu, Q., Abou-Mansour, E., L'Haridon, F., Métraux, J., Gruissem, W., Hofius, D. and Hennig, L. (2015) Chromatin assembly factor CAF-1 represses priming of plant defence response genes. *Nature Plants* **1**, 15127.

Muchova, V., Amiard, S., Mozgova, I., Dvorackova, M., Gallego, M.E., White, C. and Fajkus, J. (2015) Homology-dependent repair is involved in 45S rDNA loss in plant CAF-1 mutants. *Plant J* **81**, 198-209.

Murata, M., Ogura, Y. and Motoyoshi, F. (1994) Centromeric repetitive sequences in Arabidopsis thaliana. *Jpn J Genet* **69**, 361-370.

Nabatiyan, A. and Krude, T. (2004) Silencing of Chromatin Assembly Factor-1 in human cells leads to cell death and loss of chromatin assembly during DNA synthesis. *Mol Cell Biol* **24**, 2853-2862.

Ono, T., Kaya, H., Takeda, S., Abe, M., Ogawa, Y., Kato, M., Kakutani, T., Scheid, O.M., Araki, T. and Shibahara, K. (2006) Chromatin Assembly Factor-1 ensures the stable maintenance of silent chromatin states in Arabidopsis. *Genes Cells* **11**, 153-162.

Pavlistova, V., Dvorackova, M., Jez, M., Mozgova, I., Mokros, P. and Fajkus, J. (2016) Phenotypic reversion in fas mutants of Arabidopsis thaliana by reintroduction of FAS genes: variable recovery of telomeres with major spatial rearrangements and transcriptional reprogramming of 45S rDNA genes. *Plant J*.

Ramirez-Parra, E. and Gutierrez, C. (2007a) E2F regulates *FASCIATA1*, a chromatin assembly gene whose loss switches on the endocycle and activates gene expression by changing the epigenetic status. *Plant Physiol* **144**, 105-120.

Ramirez-Parra, E. and Gutierrez, C. (2007b) The many faces of Chromatin Assembly Factor 1. *Trends Plant Sci* **12**, 570-576.

Ray-Gallet, D., Woolfe, A., Vassias, I., Pellentz, C., Lacoste, N., Puri, A., Schultz, D.C., Pchelintsev, N.A., Adams, P.D., Jansen, L.E. and Almouzni, G. (2011) Dynamics of histone H3 deposition in vivo reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. *Mol Cell* **44**, 928-941.

Rehrauer, H., Aquino, C., Gruissem, W., Henz, S.R., Hilson, P., Laubinger, S., Naouar, N., Patrignani, A., Rombauts, S., Shu, H., Van de, P.e.Y., Vuylsteke, M., Weigel, D., Zeller, G. and Hennig, L. (2010) AGRONOMICS1: A new resource for Arabidopsis transcriptome profiling. *Plant Physiol* **152**, 487-499.

Reinholz, E. (1966) Radiation induced mutants showing changed inflorescence characteristics. *Arabid Inf Serv* **3**, 19-20.

Rutowicz, K., Puzio, M., Halibart-Puzio, J., Lirski, M., Kotlinski, M., Kroten, M.A., Knizewski, L., Lange, B., Muszewska, A., Sniegowska-Swierk, K., Koscielniak, J., Iwanicka-Nowicka, R., Buza, K., Janowiak, F., Zmuda, K., Joesaar, I. and Laskowska-Ka (2015) A Specialized Histone H1 Variant Is Required for Adaptive Responses to Complex Abiotic Stress and Related DNA Methylation in Arabidopsis. *Plant Physiol* **169**, 2080-2101.

Schönrock, N., Exner, V., Probst, A., Gruissem, W. and Hennig, L. (2006) Functional genomic analysis of CAF-1 mutants in *Arabidopsis thaliana*. *J. Biol. Chem.* **281**, 9560-9568.

Schwartz, U. and Langst, G. (2016) Bioinformatic Analysis of ChIP-seq Data on the Repetitive Ribosomal RNA Gene. *Methods Mol Biol* **1455**, 225-230.

Shibahara, K. and Stillman, B. (1999) Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* **96**, 575-585.

Shu, H., Gruissem, W. and Hennig, L. (2013) Measuring Arabidopsis chromatin accessibility using DNase I-polymerase chain reaction and DNase I-chip assays. *Plant Physiol* **162**, 1794-1801.

Shu, H., Nakamura, M., Siretskiy, A., Borghi, L., Moraes, I., Wildhaber, T., Gruissem, W. and Hennig, L. (2014) Arabidopsis replacement histone variant H3.3 occupies promoters of regulated genes. *Genome Biol* **15**.

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

Smith, S. and Stillman, B. (1989) Purification and characterization of CAF-1, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell* **58**, 15-25.

Song, Y., He, F., Xie, G., Guo, X., Xu, Y., Chen, Y., Liang, X., Stagljar, I., Egli, D., Ma, J. and Jiao, R. (2007) CAF-1 is essential for Drosophila development

and involved in the maintenance of epigenetic memory. *Dev Biol* **311**, 213-222.

Stillman, B. (1986) Chromatin assembly during sv40 DNA replication in vitro. *Cell* **45**, 555-565.

Storey, J.D. and Tibshirani, R. (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* **100**, 9440-9445.

Stroud, H., Greenberg, M.V., Feng, S., Bernatavichute, Y.V. and Jacobsen, S.E. (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell* **152**, 352-364.

Tagami, H., Ray-Gallet, D., Almouzni, G. and Nakatani, Y. (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**, 51-61.

Talbert, P.B. and Henikoff, S. (2010) Histone variants--ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* **11**, 264-275.

Winkler, D.D., Zhou, H., Dar, M.A., Zhang, Z. and Luger, K. (2012) Yeast CAF-1 assembles histone (H3-H4)2 tetramers prior to DNA deposition. *Nucleic Acids Res* **40**, 10139-10149.

Xi, Y., Yao, J., Chen, R., Li, W. and He, X. (2011) Nucleosome fragility reveals novel functional states of chromatin and poises genes for activation. *Genome Res* **21**, 718-724.

Yu, Z., Liu, J., Deng, W.M. and Jiao, R. (2015) Histone chaperone CAF-1: essential roles in multi-cellular organism development. *Cell Mol Life Sci* **72**, 327-337.

Zhang, T., Zhang, W. and Jiang, J. (2015) Genome-wide nucleosome occupancy and positioning and their impact on gene expression and evolution in plants. *Plant Physiol* **168**, 1406-1416.

Zhang, X., Bernatavichute, Y.V., Cokus, S., Pellegrini, M. and Jacobsen, S.E. (2009) Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol* **10**, R62.

Tables

Table 1. Genome regions with reduced sequence coverage in *fas2*. Sequence coverage in Col and *fas2* was compared for bins of 1000 bp. Regions (single bins or merged adjacent significant bins) with a significantly (one-sided binomial test, multiple testing correction according to (Storey et al., 2003), q<0.05) reduced coverage in *fas2* are listed.

Chromosome	Start	End	Counts in Col	Counts in *fas2*	log2FC	Comment
Chr2	1,001	11,000	188,137	22,542	-3	NOR2 border region[1]
Chr2	7,191,001	7,192,000	5,551	625	-3	*AT2G16586*
Chr3	14,191,001	14,206,000	3,069	449	-3	Pericentromeric region[2]
Chr5	3,253,001	3,254,000	16,699	1,833	-3	Upstream[3] of *AT5G10340*

[1]Includes: *AT2G01008*, *AT2G01010* (18S rRNA), *AT2G01020* (5.8S rRNA), *AT2G01021*, *AT2G01023*, *AT2G01022*

[2]Includes: *AT3G41761*, *AT3G41762*, *AT3G41768* (18S rRNA), *AT3G41979* (5.8S rRNA)

[3]Upstream: within 1000 bp from the 5´end of the gene

Table 2. Genome distribution of nucleosomes that are depleted in *fas2*.

	No. of nucleosomes in Col	%[1]	No. of depleted nucleosomes	%[2]	FC[3]	p-value (reduced freq.)	p-value (increased freq.)
Bodies of active genes	255,563	45.4	46,800	36.5	0.8	0	1
Promoters of active genes	13,884	2.5	6,957	5.4	2.17	1	0
Bodies of inactive genes	55,061	9.8	9,433	7.4	0.75	2.00E-255	1
Promoters of inactive genes	4,875	0.9	1,603	1.3	1.39	1	1.37E-59
Bodies of TE genes	36,323	6.4	6,536	5.1	0.8	7.84E-117	1
Intergenic regions	197,645	35.1	56,868	44.4	1.26	1	0

[1]Percentage of all detected nucleosomes

[2]Percentage of all depleted nucleosomes

[3]Fold change of Percentage of all depleted nucleosomes (i.e. observed) to Percentage of all detected nucleosomes (i.e. expected)

Table 3. Enriched GO terms in genes with the 10% most depleted nucleosomes in their 1 kb promoters (-1000:TSS) or with the 20% most depleted nucleosomes their 200 bp proximal promoters (-200:TSS).

GO term	q-value
1 kb promoters[1]	
Response to chitin[2]	4.90E-05
Plant-type hypersensitive response[2]	1.56E-03
Hormone-mediated signaling pathway	1.63E-03
Jasmonic acid mediated signaling pathway	4.11E-03
Negative regulation of defense response[2]	6.96E-03
Regulation of plant-type hypersensitive response[2]	7.00E-03
Salicylic acid mediated signaling pathway[2]	7.53E-03
Signal transduction	9.29E-03
Cell death	9.58E-03
Response to jasmonic acid	1.08E-02
Innate immune response[2]	1.13E-02
MAPK cascade	1.27E-02
Regulation of innate immune response[2]	1.38E-02
Response to salicylic acid[2]	1.70E-02
Defense response[2]	2.41E-02
Response to endoplasmic reticulum stress	3.06E-02
Response to fungus[2]	3.88E-02
Cellular response to abscisic acid stimulus	4.69E-02
Regulation of defense response[2]	4.69E-02
200 bp proximal promoters[1]	
Response to chitin[2]	6.30E-08
Innate immune response[2]	7.70E-05
Defense response[2]	2.20E-04
Plant-type hypersensitive response[2]	5.00E-04
Respiratory burst involved in defense response[2]	9.80E-04
Hyperosmotic response	1.90E-03
Regulation of innate immune response[2]	3.60E-03
Defense response, incompatible interaction[2]	4.20E-03
Response to fungus[2]	4.40E-03
Response to other organism	5.30E-03
Regulation of plant-type hypersensitive response[2]	6.40E-03
Defense response to fungus[2]	8.90E-03
Salicylic acid mediated signaling pathway[2]	1.10E-02
Cell death	1.40E-02
Response to salicylic acid[2]	1.60E-02
Regulation of defense response[2]	3.20E-02

[1]Most depleted nucleosomes are those nucleosomes from the set of significantly depleted nucleosomes that have the largest fold change in occupancy.

[2]GO terms directly related to plant defense.

Table 4. Genome distribution of nucleosomes that are shifted in *fas2*.

	No. of nucleosomes in Col	%[1]	No. of shifted nucleosomes	%[2]	FC[3]	p-value (reduced freq.)	p-value (increased freq.)
Bodies of active genes	255,563	45.4	12,628	48.8	1.08	1	1.48E-30
Promoters of active genes	13,884	2.5	303	1.2	0.47	2.72E-52	1
Bodies of inactive genes	55,061	9.8	2,945	11.4	1.16	1	9.84E-19
Promoters of inactive genes	4,875	0.9	148	0.6	0.64	2.04E-08	1
Bodies of TE genes	36,323	6.4	2,469	9.5	1.49	1	1.92E-85
Intergenic regions	197,645	35.1	7,371	28.5	0.81	4.55E-118	1

[1]Percentage of all detected nucleosomes

[2]Percentage of all shifted nucleosomes

[3]Fold change of Percentage of all shifted nucleosomes (i.e. observed) to Percentage of all detected nucleosomes (i.e. expected)

Table 5. Enriched GO terms in genes with the 10% most shifted nucleosomes in their gene bodies.

GO term	q-value
Transmembrane transport	2.42E-03
Reproductive structure development[1]	2.81E-03
Multicellular organismal development[1]	2.91E-03
Shoot system development[1]	3.41E-03
Tissue development[1]	1.70E-02
Flower development[1]	2.60E-02
Anatomical structure morphogenesis[1]	2.76E-02
Auxin efflux	3.65E-02
Meristem maintenance[1]	3.83E-02
Regulation of cell proliferation	3.84E-02

[1]GO terms directly related to plant development

Figure Legends

Figure 1. Confirmation of loss of genome regions.

Quantitative PCR-based verification of genomic DNA depletion identified by DNA genome-wide resequencing (NGS). (A) Single-copy regions used as non-depleted controls – *ACT2* (*AT3G18780*) and *PP2A* (*AT1G13320*). First, the amount of *ACT2* amplicon was normalized to wild-type (WT) level to confirm comparable amount of the *ACT2* DNA in *fas2* and WT. All the remaining results were normalized to *ACT2* levels. (B) 18S rDNA-specific primers confirm the expected depletion of 45S rDNA in *fas2*. (C) Verification of DNA depletion in the 4 genomic regions in *fas2* identified by NGS. Regions that were found to be repetitive respective to *ACT2* were used as queries in a BLAST (NCBI) search against the *Arabidopsis thaliana* nucleotide collection (nr/nt) and identified as 45S rDNA or 5S rDNA sequences as indicated by the blue label. Note that the depletion of the *fas2* genomic DNA region on chromosome 5 identified by NGS was not confirmed by qPCR. (D) Genomic regions bordering 45S rDNA cluster on chromosome 4 (NOR4) are not affected in *fas2*. Red arrowheads indicate the location of regions amplified by qPCR.

Figure 2. Localization of depleted nucleosomes. Nucleosomes around the TSS of all genes were grouped in 100 bp bins and significance of the frequency of depleted nucleosomes was calculated for each bin using hypergeometric tests followed by multiple test correction according to Bonferroni. (A) All genes, (B) genes without expression changes in *fas2*, (C) genes with expression changes in *fas2*. Horizontal dashed lines indicate $-\log_{10}$(p-value) equal to +10 and -10, vertical solid line indicate TSS, vertical dashed lines indicate +/- 200 bp relative to the TSS.

Figure 3. Localization of shifted nucleosomes. Nucleosomes around the TSS of all genes were grouped in 100 bp bins and significance of enrichment of shifted nucleosomes was calculated for each bin using hypergeometric tests followed by multiple test correction according to Bonferroni. Tests were performed for all shifted nucleosomes (blue lines), nucleosomes shifted forward (green lines) or reverse (purple lines) relative to the direction of transcription. (A) All genes and (B) genes without expression changes in *fas2*. Horizontal dashed lines indicate $-\log_{10}$(p-value) equal to +10 and -10, vertical solid line indicate TSS, vertical dashed lines indicate +/- 200 bp relative to the TSS.

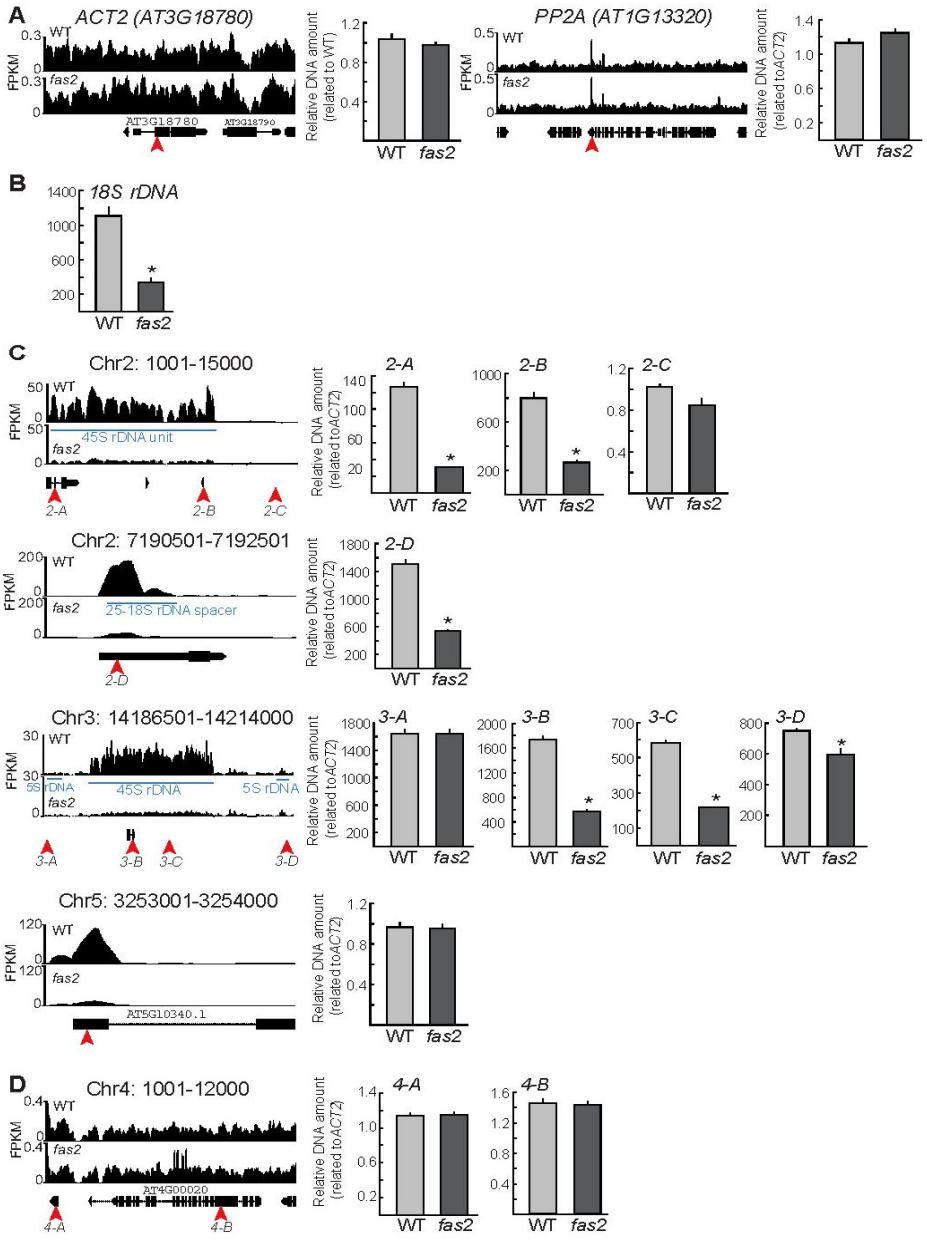

Figure 1. Confirmation of loss of genome regions.
Quantitative PCR-based verification of genomic DNA depletion identified by DNA genome-wide resequencing (NGS). (A) Single-copy regions used as non-depleted controls – *ACT2* (*AT3G18780*) and *PP2A* (*AT1G13320*). First, the amount of *ACT2* amplicon was normalized to wild-type (WT) level to confirm comparable amount of the *ACT2* DNA in *fas2* and WT. All the remaining results were normalized to *ACT2* levels. (B) 18S rDNA-specific primers confirm the expected depletion of 45S rDNA in *fas2*. (C) Verification of DNA depletion in the 4 genomic regions in *fas2* identified by NGS. Regions that were found to be repetitive respective to *ACT2* were used as queries in a BLAST (NCBI) search against the *Arabidopsis thaliana* nucleotide collection (nr/nt) and identified as 45S rDNA or 5S rDNA sequences as indicated by the blue label. Note that the depletion of the *fas2* genomic DNA region on chromosome 5 identified by NGS was not confirmed by qPCR. (D) Genomic regions bordering 45S rDNA cluster on chromosome 4 (NOR4) are not affected in *fas2*. Red arrowheads indicate the location of regions amplified by qPCR.

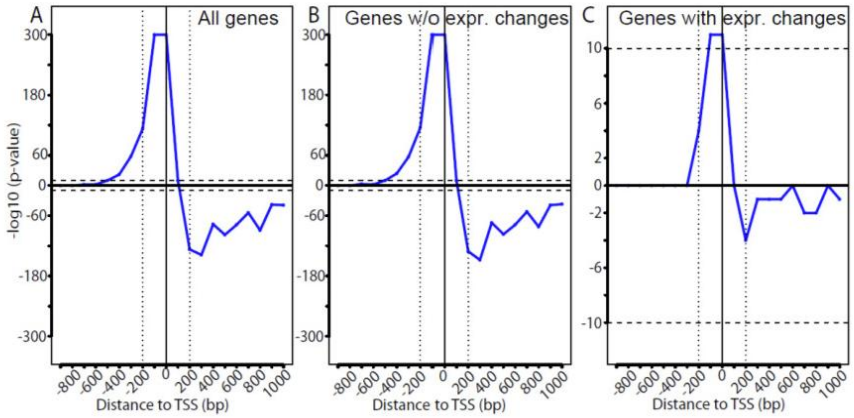

Figure 2. Localization of depleted nucleosomes. Nucleosomes around the TSS of all genes were grouped in 100 bp bins and significance of enrichment of depleted nucleosomes was calculated for each bin using hypergeometric tests followed by multiple test correction according to Bonferroni. (A) All genes, (B) genes without expression changes in *fas2*, (C) genes with expression changes in *fas2*. Horizontal dashed lines indicate $-\log_{10}$(p-value) equal to +10 and -10, vertical solid line indicate TSS, vertical dashed lines indicate +/- 200 bp relative to the TSS.

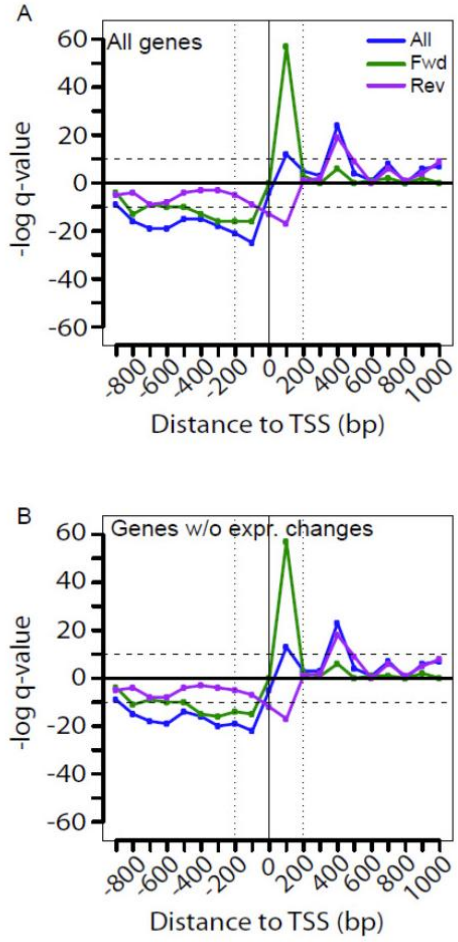

Figure 3. Localization of shifted nucleosomes. Nucleosomes around the TSS of all genes were grouped in 100 bp bins and significance of enrichment of shifted nucleosomes was calculated for each bin using hypergeometric tests followed by multiple test correction according to Bonferroni. Tests were performed for all shifted nucleosomes (blue lines), nucleosomes shifted forward (green lines) or reverse (purple lines) relative to the direction of transcription. (A) All genes and (B) genes without expression changes in *fas2*. Horizontal dashed lines indicate $-\log_{10}$(p-value) equal to +10 and -10, vertical solid line indicate TSS, vertical dashed lines indicate +/- 200 bp relative to the TSS.