Title
Chromatin Structure-Mediated Regulation of Nuclear Processes

Author
Kim, Min

Publication Date
2013

Peer reviewed|Thesis/dissertation

Chromatin Structure-Mediated Regulation of Nuclear Processes

by

Min Kim

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Plant Biology

in the

Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Daniel Zilberman, Chair
Professor Robert Fischer
Professor Rebecca Heald
Professor Jasper Rine

Fall 2013

Chromatin Structure-Mediated Regulation of Nuclear Processes

Abstract

Chromatin Structure-Mediated Regulation of Nuclear Processes

by

Min Kim

Doctor of Philosophy in Plant Biology

University of California, Berkeley

Professor Daniel Zilberman, Chair

Chromatin is a mixture of DNA and DNA binding proteins that control transcription. Dynamic chromatin structure modulates gene expression and is responsible for an extraordinary spectrum of developmental processes. An intricate interplay of DNA methylation, histone modifications, histone variants, small RNA accumulation, and ATPase chromatin remodelers defines chromatin re-configuration in a precise manner, locally within a cell and globally across different cell types.

The development of high-throughput screening methods such as microarray and whole-genome sequencing has led to an explosion of chromatin studies in the past decade. Moreover, genetic and molecular studies resulted in identification of a number of proteins that may influence chromatin structure. However, the exact functions of individual proteins as well as their functional relationships to each other are less understood. Also, the role of chromatin components in establishing cell- and tissue-specific chromatin structure is largely unknown. To address these open questions in chromatin biology, I focused my dissertation work on 1) studying tissue-specific DNA demethylation in seed, and 2) determining the role of a ubiquitous DNA binding protein, linker histone H1, in regulating chromatin structure.

Tissue specific DNA methylation in seed. In endosperm, the nutritive tissue that nourishes the embryo, parent-of-origin specific gene expression is regulated by DNA demethylation. However, the extent to which DNA demethylation occurs in a tissue-specific manner and regulates transcription in the endosperm of crop plants like rice remains unknown. To address these questions, my colleagues and I examined the DNA methylation patterns of two rice seed tissues, embryo and endosperm. We found that endosperm genome is globally hypomethylated at non-CG sites and locally hypomethylated at CG-sites compared to embryo. We also identified that

small transposons near genes (euchromatic regions) are the primary targets of DNA demethylation. The loci near the genes preferentially expressed in endosperm (e.g. storage protein and starch synthesizing enzymes) are subjected to local hypomethylation, suggesting that DNA methylation plays a role in inducing tissue-specific genes in endosperm.

The role of H1 in regulating chromatin structure. H1 is proposed to facilitate higher order chromatin structure, but its effects on individual chromatin components and transcription are less understood. To resolve this issue, we investigated the role of H1 in regulating DNA methylation, nucleosome positioning, and transcription. We identified that H1 was most enriched in transposons. H1 was also found in genes at a lower level compared to transposons, and the abundance of H1 was anticorrelated with gene expression. Moreover, H1 influences nucleosome positioning by increasing the distance between two nucleosomes. Lack of H1 resulted in increased DNA methylation of transposons with heterochromatic features. In contrast, an *h1* mutant showed a reduction of DNA methylation in genes and transposons with euchromatic features. Our finding suggests that H1 has a dual function in regulating DNA methylation. That is, H1 inhibits both DNA methyltransferases and DNA demethylation-associated enzymes from binding heterochromatin and euchromatin, respectively. In addition, the hypermethylated loci in our *h1* mutant almost perfectly overlapped with the hypomethylated loci in a *ddm1* mutant in heterochromatin, suggesting a link between these two proteins. DDM1 is an Snf2 chromatin remodeler that can slide nucleosomes along DNA and has been proposed to provide DNA methyltransferase access to target sequences. We further determined their functional relationship by crossing *h1* and *ddm1* mutants, and generated a map of DNA methylation of the cross. We identified that loss of DNA methylation from *ddm1* was partially recovered by removing H1. Also the mutant phenotype observed in *ddm1* disappeared in *h1ddm1*. Based on our results, we proposed a model where DDM1-mediated chromatin destabilization releases H1 binding, which in turn increases DNA accessibility.

It is noteworthy that DNA demethylation preferentially occurred in euchromatin in both the rice seed DNA methylation study and the H1 study. Based on this result, we proposed that the apparent target preference of DNA demethylation-associated proteins depends on the underlying chromatin structure. We think that this chromatin structure-mediated specificity also dictates other nucleoproteins to determine/recognize their targets.

My dissertation work tackled multiple aspects of chromatin biology: tissue-specific chromatin regulation, and the interplay between chromatin components in chromatin organization. Together, the results from my work enhanced our knowledge of how chromatin components influence overall chromatin structure.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my advisor Dr. Daniel Zilberman and my thesis committee members, Dr. Robert Fischer, Dr. Jasper Rine, and Dr. Rebecca Heald for all their advices and support throughout graduate school. I would also like to thank the Zilberman lab and the Fischer lab members for the stimulating and enjoyable academic environment; and Dr. Takashi Okamoto and his lab members for the awesome research experience in Japan. I also thank Dr. Renee Sung, who introduce me to science, for her continuous support.

I especially thank Dr. Young-wook Jun (more known as Dr. Wa) who has been my 101 guide on how to survive academic life. Lastly, many, many thanks to my wonderful friends and family, I would not have survived graduate school without you.

Chapter 1:

Introduction

Highly complex multicellular organisms originate from a single cell. This single cell propagates into individual cells with different cell fates, and the interplay of these differentiated cells generates structures such as tissues and organs that differ in shape and function. The underlying DNA sequences of differentiated cells are identical; thus, genetic material change is not responsible for the difference. Instead, a characteristic spatial and temporal array of protein-coding gene expression is observed in cells with a specific fate. The unique transcriptional program of a cell depends on the cell-specific cocktail of transcription factors induced by developmental or environmental signals. These transcription factors then recruit RNA Polymerase II (Pol II) to target loci. The precise regulation of transcription factors is crucial in cell differentiation because ectopic expression of a few key transcription factors can reprogram and switch one cell fate to another, as shown in induced pluripotent stem cells (Yamanaka, 2012).

The binding of transcriptional machinery to specific targets is modulated by the unique packaging of the genome in chromatin (Woodcock and Ghosh, 2010). Chromatin controls gene expression, and thus defines a wide range of biological processes in plants and animals (Li and Reinberg, 2011). This process involves a dynamic interplay between multiple key components: nucleosomes, DNA- and histone-modifying proteins, and chromatin-remodeling proteins (Li and Reinberg, 2011). These components alter protein-DNA interaction, thus changing the overall chromatin structure, DNA accessibility, and in turn transcription (Kornberg and Lorch, 1992). Moreover, a function for chromatin structure has been implicated in the more general and global regulation of all DNA binding proteins, which attests to the importance of chromatin structural regulation beyond transcriptional regulation (Li and Reinberg, 2011; Woodcock and Ghosh, 2010). There have been many new reports on chromatin-mediated transcriptional regulation, but it still remains unclear how chromatin components specifically target certain silent or active regions of the chromatin to finally regulate transcription. In this chapter, as an introduction, I begin with describing the general organization of chromatin, followed by a detailed description of the known mechanisms of methylation and demethylation of plant DNA in detail. Recent studies on chromatin make it apparent that dynamic regulation of chromatin impacts transcription and development (Law and Jacobsen, 2010; Li and Reinberg, 2011; Zentner and Henikoff, 2013).

1.1 Chromatin structure

1.1.1 Nucleosome

Nucleosomes are the elementary repeating subunit of chromatin, appearing as "beads on a string" under electron microscopy, with beads representing nucleosomes and the string as the linker DNA between two nucleosomes (Olins and Olins, 1974; Woodcock et al., 1976). 75-90% of genomic DNA is presumably wrapped in nucleosomes (Tolkunov and Morozov, 2009), emphasizing the importance of nucleosomes in chromatin structure organization. The nucleosome consists of a core histone octamer (i.e. two sets of histones H2A, H2B, H3 and H4), wrapped by approximately 150 bp of DNA in left-handed superhelical turns (Li et al., 2007a). The proposed model for nucleosome assembly occurs in a step-wise manner, starting with two H3-H4 dimers forming a tetramer which is deposited onto DNA. Next, the two sets of H2A-H2B dimers join the tetramer-DNA structure to finalize the assembly of a core octamer (Akey and Luger, 2003; Burgess and Zhang, 2013; Luger et al., 1997). These histones are highly basic due to a high lysine and arginine amino acids content (Arents and Moudrianakis, 1993). The positively charged histones form strong electrostatic bonds with negatively charged DNA, making nucleosome-associated nucleotides less accessible to DNA factors such as transcriptional machinery than nucleosome-free DNA (Bell et al., 2011). Thus, precise nucleosome positioning is important in regulating gene expression. Transcriptionally active genes show a depletion of nucleosomes at proximal promoters upstream of the transcription start sites (TSS), which need to be accessed by transcriptional machinery (Bai et al., 2010; Bell et al., 2011). This notion is supported by the high DNAse-I accessibility at nucleosome-depleted regions (Kodama et al., 2007; Spiker et al., 1983; Zhang et al., 2012). Consistently, disruption of nucleosome positioning at the TSS results in production of cryptic transcripts (Whitehouse et al., 2007). During nuclear processes such as DNA replication, gene transcription, and DNA repair, which require enhanced genome accessibility, the nucleosomes are dynamically reorganized (Eaton et al., 2010; 2011; Kornberg and Lorch, 1992; Schwaiger et al., 2009; Thoma, 2005; van Attikum and Gasser, 2005). This process is determined by various factors including intrinsic DNA sequences, DNA binding proteins, and nucleosome remodelers.

Nucleosome properties mediated by histone modification

Histone modification impacts transcriptional activation and inhibition. Histone domains, which include a long N-terminus, a C-terminus, and a globular domain, are extensively acetylated, methylated, phosphorylated, and ubiquitinated (Loidl, 2004; Margueron et al., 2005). These modifications alter the histone-DNA interaction and recruit chromatin remodelers, impacting overall chromatin structure for transcriptional regulation. The exact function of most histone modifications is unclear, but some are thought to act through charge neutralization of chromatin, creating more open or closed chromatin configurations (Schones and

Zhao, 2008) and facilitating the binding of proteins that recruit transcriptional machinery (Liu et al., 2010). In particular, histone acetylation and methylation are key modifications for regulating DNA accessibility and hence levels of gene transcription.

Histone acetylation neutralizes positive charges on histones, thus weakening the interactions between histones and the negative charged DNA phosphate backbone, as well as the interactions between nucleosomes and interacting proteins. This produces a more relaxed chromatin structure and increases DNA accessibility (McGhee and Felsenfeld, 1980). Histone acetylation is catalyzed by histone acetyltransferases (HATs) using acetyl-CoA as a substrate (Shahbazian and Grunstein, 2007). HATs generally have low substrate specificity, and acetylate multiple lysines on histone tails (Loidl, 2004; Margueron et al., 2005). The acetylation modification can be reversed by histone deacetylases, and removal of acetylation is correlated with a repression of transcription (Loidl, 2004; Zentner and Henikoff, 2013). The genome-wide map of H3 lysine 9 acetylation (H3K9ac) in *Arabidopsis* showed accumulation of H3K9ac preferentially at the TSS (Zhou et al., 2010). The enrichment is positively correlated with transcription levels, indicating that histone acetylation may have a function in facilitating transcription initiation and elongation (Nightingale et al., 1998).

Histone methylation is catalyzed by histone methyltransferases (HMTs) (Liu et al., 2010). HMTs can add up to three methyl marks on a lysine residue, creating mono-methyl lysine, di-methyl lysine, and tri-methyl lysine (Liu et al., 2010). Activation or repression of transcription depends on the methylation of specific lysine residues (e.g, K9, K27, K4, etc) and the level of methylation on the lysine (Liu et al., 2010; Martin and Zhang, 2005). For example, histone di-methylation on lysine 9 (H3K9me2) is associated with heterochromatin formation in a wide range of eukaryotic species (Mathieu et al., 2005), H3K27me1/2/3 are associated with transcription silencing (Jacob et al., 2010), and H3K4 methylation with gene activation (Calo and Wysocka, 2013).

Unlike acetylation, methylation does not change the overall charge of histones. Therefore, the mechanism by which histone methylation controls transcription may be less direct compared to histone acetylation. Histone methylation may function as a binding site for proteins necessary for chromatin reconfiguration. For example, Heterochromatin Protein 1 (HP1), a component of heterochromatin, binds to H3K9me2 in *Drosophila* (Bannister et al., 2001). The *Arabidopsis* homolog of HP1, Like-HP1 (LPH1), regulates gene silencing (Mylne et al., 2006) and associates with H3K27me3 (Turck et al., 2007; Zhang et al., 2007), suggesting that LHP1 binds H3K27me3. Also, CMT3, a chromodomain-containing DNA methyltransferase in *Arabidopsis*, binds to H3K9me2, creating a more compact heterochromatin structure (Du et al., 2012; Jackson et al., 2002). Furthermore, PHD containing proteins such as the large subunit of the origin recognition complex, ORC1, ING, and Alfin1-like are identified to interact with H3K4me in *Arabidopsis* (la Paz

Sanchez and Gutierrez, 2009; Lee et al., 2009). Lastly, the mammalian H3K4me binding protein CHD1 is an ATPase chromatin remodeler and a part of a histone acetyltransferase complex (Pray-Grant et al., 2005; Wysocka et al., 2006), suggesting that H3K4me acts in part as a signal to recruit a chromatin remodeler and reinforces gene activation.

Sequence-dependent nucleosome positioning

Nucleosome formation occurs preferentially at certain DNA sequences (Bell et al., 2011; Struhl and Segal, 2013). First, the periodicity of certain nucleotides has been implicated in the preferential formation of nucleosomes. The periodic dinucleotide, AT and TA, is associated with stable nucleosome formation (Segal et al., 2006; Takasuka and Stein, 2010). This AT/TA dinucleotide is preferentially found in the minor grooves of DNA and the crystal structure of a nucleosome shows that the minor grooves directly interact with histones (Suto et al., 2003). Nucleosome positioning is also associated with a stretch of poly(dA:dT) (Anderson and Widom, 2001). The stretches of A/Ts are intrinsically stiff and may inhibit nucleosome binding at that site. Promoters with a TATA box (Molina and Grotewold, 2005; Yang et al., 2007) are depleted of nucleosomes and are a general binding site for transcriptional factors and Pol II.

Distributions of nucleosomes *in vitro* and *in vivo* for the same genome are similar (Kaplan et al., 2009), indicating that intrinsic sequence is important for nucleosome positioning. However, there are important differences between the two samples and among *in vivo* samples with different growth conditions (Kaplan et al., 2009), indicating that there are other mechanisms to reposition nucleosomes. Also, it has been identified that replication and DNA repair involve dramatic re-organization of chromatin (Eaton et al., 2010; Schwaiger et al., 2009; Thoma, 2005; Ura et al., 2001; van Attikum and Gasser, 2005), which likely requires active nucleosome remodeling mechanisms to increase or inhibit DNA accessibility. The known mechanism to alter nucleosome positioning involves ATPase chromatin remodelers which will be discussed in the following section.

Dynamic nucleosome positioning via Snf2 ATPase chromatin remodelers

Snf2 helicase-related ATPases form a large family of proteins first discovered in yeast (Winston and Carlson, 1992) and later identified to be conserved in eukaryotic organisms (Clapier and Cairns, 2009; Knizewski et al., 2008). The proteins have a functional ATPase domain but the helicase domain of Snf2 protein is not functional,

as the proteins cannot separate double stranded DNA (Dürr et al., 2006). They have a 3'->5' translocase activity on naked DNA, which generates a torsional strain on DNA (Bowman, 2010). This strain presumably perturbs histone-DNA contact *in vivo* and allows the core octamers to move. This ATP-dependent movement of nucleosomes results in nucleosome sliding, ejection, and unwrapping of histone octamers ("repositioning") as well as the exchange of histone variants and removal of certain histones ("reassembling") (Clapier and Cairns, 2009; Ryan and Owen-Hughes, 2011). The "repositioning" mechanism allows the exposure of target sequences to proteins. A Snf2 protein was shown to be involved in expression of an inducible transcript, indicating that remodelers allow transcription machinery to bind to otherwise hidden and inactive genes upon an external or internal signal (Schwabish and Struhl, 2007). The "reassembling" mechanism is utilized in eviction of H2A-H2B dimers. H2A.Z, a H2A variant, is exchanged with the canonical H2A via Swr1-like Snf2 ATPase (Luk et al., 2010). Another Snf2 protein, INO80, is known to evict H2A.Z (Papamichos-Chronakis et al., 2011), indicating that the H2A.Z exchange is dynamic and reversible. H2A.Z deposition is enriched in inducible, responsive genes (Coleman-Derr and Zilberman, 2012), suggesting that the role of a Snf2 remodeler in exchanging histones is linked with transcriptional regulation.

Chromatin remodelers play a crucial role in almost every known nuclear processes, such as DNA methylation (Chan et al., 2006; Kanno et al., 2004; Zemach et al., 2013), histone modification (Pray-Grant et al., 2005), DNA repair (Rosa et al., 2013), and replication (Vincent et al., 2008), emphasizing the importance of the Snf2 family proteins. The exact molecular function, however, is less understood. Therefore, identifying key chromatin remodelers and determining their exact function are current topics of interest.

1.1.2 Higher order chromatin structure

Higher order chromatin structure—higher degrees of nucleosome compaction—impacts DNA accessibility and transcription (Woodcock and Ghosh, 2010). Chromatin compaction beyond nucleosome formation depends on physiological salt concentration, histone modifications (e.g. acetylation), histone variants (e.g. H2A.Z), chromatin-binding proteins (e.g. HP1 and LHP1), high mobility proteins, and histone H1 (Arya and Schlick, 2009; Li and Reinberg, 2011). Specifically, in this section, I focus on H1-mediated chromatin structure.

H1-mediated chromatin compaction

H1 is as abundant as core histones and highly conserved throughout eukaryotic organisms (Bates and Thomas, 1981; Kasinsky et al., 2001). H1 has less conserved N- and C-termini than core histones, but has a greatly conserved globular domain (Kasinsky et al., 2001; Wierzbicki and Jerzmanowski, 2005). The tails are highly positively charged due to a high lysine content, a common characteristic of H1 in eukaryotic organisms (Kasinsky et al., 2001). Unlike core histones, H1 is not a part of the nucleosomes, but interacts with nucleosomes and the linker DNA of ~20 bp (Ramakrishnan et al., 1993; Zhou et al., 1998). H1 binds to the dyad axis and entrance/exit site of DNA in a nucleosome (Ramakrishnan et al., 1993), and the location of H1 suggests that it acts as a node, restricting nucleosome movement (Fan and Roberts, 2006; Pennings et al., 1994). Its lysine-rich tails likely bind to the negatively charged linker DNA, and facilitate higher order chromatin condensation (Lu and Hansen, 2004; Misteli et al., 2000; Routh et al., 2008). While H1 is abundant in heterochromatin, the compacted region of chromatin, it is less abundant in euchromatin, the gene-rich transcriptionally-active region of chromatin (Ascenzi and Gantt, 1999; Bresnick et al., 1992; Lever et al., 2000; Weintraub, 1984).

Initially, it was thought that H1 statically binds to a nucleosome to regulate transcription. However, fluorescence recovery after photobleaching (FRAP) experiments with GFP-tagged H1 in a live nucleus revealed that H1 is highly mobile and H1-nucleosome interaction is rather transient (Lever et al., 2000; Misteli et al., 2000). The transient interaction may be influenced by H1 phosphorylation, which leads to decreased nucleosome-H1 binding stability. Phosphorylation on serine and threonine residues of the H1 tails neutralizes the charge and reduces H1 binding affinity for DNA (Garcia et al., 2004). Constitutively phosphorylated H1 mimics the loss of H1 in chromatin and creates an overall more relaxed chromatin structure (Dou and Gorovsky, 2000; Herrera et al., 1996). Also, H1 phosphorylation facilitates the transcription by RNA polymerase I and II (Zheng et al., 2010), suggesting that H1 regulates transcription.

Dynamic regulation of transcription and chromatin structure by histone H1

Loss of H1 increases DNA digestion by Micrococcal nuclease, an enzyme which preferentially digests naked DNA, indicating that low levels of H1 generate a more accessible chromatin structure (Barra et al., 2000; Belikov et al., 2007). Higher chromatin accessibility is correlated with an increase in transcription (Görisch et al., 2005; Li et al., 2007a), therefore a global change of gene expression has been anticipated from the loss of H1. Interestingly, removal of H1 or hyperphosphorylating H1 did not globally alter the transcription and impacted only a few genes (Dou and Gorovsky, 2000; Fan et al., 2005; Hellauer et al., 2001; Lu et

al., 2009; Ni et al., 2006; Shen and Gorovsky, 1996). Both activation and repression of transcription was observed from the genes with altered expression levels, implying that H1 plays dual roles in transcriptional regulation, presumably by regulating the access of transcriptional activators and repressors.

H1 also regulates several key components involved in chromatin accessibility, such as heterochromatin binding proteins, histone modification H3K9me2, and DNA methylation. For example, H1 directly interacts with HP1, a protein important for heterochromatin formation (Daujat et al., 2005; Hale et al., 2006; Nielsen et al., 2001). Reduction of H1 results in the dispersion of HP1 binding patterns in *Drosophila* (Lu et al., 2009). Also, H3K9me2 virtually disappears upon removal of H1 in *Drosophila* (Lu et al., 2009). A similar reduction of H3K9me2 was observed in humans (Li et al., 2012), indicating that H1 is necessary for recruiting H3K9 methyltransferase machinery to cognate loci in eukaryotic organisms.

In addition, H1 is highly linked to DNA methylation (Rupp and Becker, 2005), a key modification associated with transcriptional silencing. Interestingly, the reduction of H1 resulted in inconsistent changes in , in various species: local hypomethylation in mammals, global hypermethylation in fungi, and local hyper- and hypo-methylation in plants (Barra et al., 2000; Fan et al., 2005; Rea et al., 2012; Wierzbicki and Jerzmanowski, 2005). The role of H1-mediated DNA methylation is still being debated and requires further investigations.

1.2 DNA methylation and demethylation in *Arabidopsis*

1.2.1 Establishment and maintenance of DNA methylation

DNA methylation is a well-studied epigenetic chromatin modification, as it plays an essential role in transcriptional silencing of transposons (Law and Jacobsen, 2010). DNA methylation is evolutionarily conserved (Zemach et al., 2010) and occurs on the fifth carbon of cytosines in CG, CHG, and CHH (where H is A, T, or C) nucleotides throughout eukaryotic organisms (Law and Jacobsen, 2010).

In particular, plant DNA methylation is predominantly found in transposons and repetitive sequences in all three cytosine contexts (Law and Jacobsen, 2010). Some genes are CG methylated and moderately expressed genes show the most methylation (Zilberman et al., 2007), suggesting that the function of genic DNA methylation differs from transposon DNA methylation. Transposons are methylated via RNA-directed DNA methylation (RdDM) and DECREASE IN DNA METHYLATION 1 (DDM1)-mediated DNA methylation pathways (Law and Jacobsen, 2010; Zemach et al., 2013).

RdDM-mediated DNA methylation establishment and maintenance

RdDM is involved in both the establishment and maintenance of DNA methylation (Law and Jacobsen, 2010; Zemach et al., 2013), and likely evolved to protect an organism against foreign viral and bacterial DNA invasion in addition to making the organism immune to transposon activation (Mathieu and Bender, 2004). RdDM uses RNAs produced from foreign sequences and transposons as "guide sequences" to identify the location of the invasive elements (Law et al., 2010; Zhang and Zhu, 2011). RdDM can be separated into two parts: RNA Polymerase IV (Pol IV)-driven RNA-induced silencing complex (RISC) assembly and Pol V-mediated sRNA target recognition and establishment of DNA methylation.

Biogenesis of 24 nucleotide (nt) sRNA initiates RdDM. Pol IV generates single stranded RNA (ssRNA), which is processed to double stranded RNA (dsRNA) by RNA-DEPENDENT RNA POLYMERASE (RDR2). RDR2 and Pol IV are co-purified and RDR2 requires Pol IV to function, suggesting that RDRD2 and Pol IV form a complex *in vivo* (Haag et al., 2012). DICER-LIKE3 (DCL3) then digests the dsRNA into 24nt sRNA (Henderson et al., 2006). Along with these key enzymes in processing sRNAs, CLASSY1 (CLSY1), a putative Snf2 ATPase chromatin remodeling protein, is proposed to act during the ssRNA and/or dsRNA processing (Smith et al., 2007). Without CLSY1, the localization of Pol IV and RDR2 is disrupted, suggesting that an ATPase chromatin remodeler is required for proper initiation of RdDM (Smith et al., 2007). Also, a putative DNA binding protein DNA-BINIDNG TRANSCRIPTION FACTOR 1 (DTF1)/SAWADEE HOMEODOMAIN HOMOLOG 1 (SHH1) has been independently identified twice as a Pol IV associated protein in a suppressor screening (Zhang et al., 2013) and in a Pol IV co-immunoprecipitation assay (Law et al., 2011). DTF1/SHH1 recognizes H3K9 methylation and binds CLSY1 (Zhang et al., 2013), suggesting that DTF1/SHH1 recruits Pol IV to genomic loci with H3K9me by interacting with CLSY1. 24nt sRNAs are loaded to ARGONAUTE 4 (AGO4) as a part of RISC for the subsequent RdDM processes (Li et al., 2006).

The second part of RdDM is Pol V-mediated target sequence recognition by sRNA. AGO4 has been identified to co-localize with Pol V and DRM2 in a discrete nuclear body, called the AB-body (Li et al., 2008). The co-localization in AB bodies implies that the AGO-siRNA complex interacts with Pol V and DRM2. Similar to CLSY1 in Pol IV pathway, an Snf2 family ATPase chromatin remodeler, DECREASE IN RNA DIRECTED DNA METHYLATION 1 (DRD1), is associated with Pol V and is absolutely required for Pol V function (Wierzbicki et al., 2008; Zemach et al., 2013). Also, Pol V is associated with a Structural Maintenance of Chromosome (SMC) domain protein, DEFECTIVE IN MERISTEM SILENCING 3 (DMS3) (Wierzbicki et al., 2009). The SMC family is another family of ATPase

chromatin remodelers involved in higher order chromatin organization (Losada and Hirano, 2005). These results suggest that DRD1 and DMS3 together induce open chromatin to recruit a Pol V complex. Once Pol V facilitates transcription of its target element, sRNA-bound AGO4 recognizes and stably associates with the Pol V complex. AGO4 has been identified to interact directly with Pol V as well as Pol V transcripts. These interactions suggest that sRNA in AGO4 binds the complementing Pol V transcript (El-Shami et al., 2007; Li et al., 2006; Wierzbicki et al., 2009). Additionally, INVOVLVED IN DE NOVO 2 (IDN2), a putative dsRNA-binding protein, is involved in the AGO4-PolV interaction, possibly stabilizing the sRNA-PolV transcript duplex (Ausin et al., 2009). Finally, a DNA methyltransferase DOMANS REARRANGED METHYLTRANSFERASE 2 (DRM2) is recruited to methylate the target DNA to facilitate transcriptional silencing.

DDM1-mediated DNA methylation maintenance

To continuously silence transposons during cell division, the DNA methylation landscape needs to be efficiently replicated in addition to the DNA sequences. While RdDM maintains DNA methylation of some repetitive sequences (Zemach et al., 2013), an alternative maintenance pathway presumably recognizes already methylated sequences and efficiently replicates the rest of DNA methylation landscape. Each cytosine context is uniquely maintained by three different DNA methyltransferases, METHYLTRANSFERASE 1 (MET1), CHROMOMETHYLASE 3 (CMT3), and CHROMOMETHYLASE 2 (CMT2) (Law and Jacobsen, 2010; Zemach et al., 2013).

DDM1, a Snf2 ATPase family chromatin remodeler, is a key component of this DNA methylation maintenance pathway (Jeddeloh et al., 1999; Lippman et al., 2004). Loss of DDM1 results in a profound reduction of DNA methylation in all cytosine contexts, suggesting that DDM1 recruits MET1, CMT3 and CMT2 to heterochromatin (Zemach et al., 2013). *In vitro*, DDM1 moves core histone octamers along DNA (Brzeski and Jerzmanowski, 2003), which implies that DDM1-mediated chromatin re-configuration increases DNA accessibility for binding of DNA methyltransferases *in vivo*.

Once DDM1 provides open access to heterochromatin, CG methylation is maintained by MET1, a plant homolog of DNA METHYLTRANSFERASE 1 (DNMT1) in mammals (Law and Jacobsen, 2010). MET1 and DNMT1 are highly similar in structure and function, thus it is generally thought that the mechanism by which MET1 methylates CG sites is similar to DNMT1 (Jurkowski and Jeltsch, 2011; Law and Jacobsen, 2010). In mammals, hemimethylated DNA is recognized by a SET- or RING-associated (SRA) domain protein that sequentially recruits DNMT1 to replication foci (Liu et al., 2013). In plants, the VARIANT IN

METHYLATION (VIM)/ORTHRUS is a family of SRA domain proteins. The SRA domain of VIM1 and VIM3 was shown to bind methylated CG sites (Johnson et al., 2007; Woo et al., 2008; 2007), suggesting that VIM proteins may function similarly to mammalian SRA domain proteins during replication. CHG methylation is maintained by CMT3, a plant specific DNA methyltransferase predominantly expressed during replication (Du et al., 2012). It has been proposed that CMT3 and H3K9 methyltransferases create a reinforcing loop to efficiently maintain CHG methylation (Jackson et al., 2002). Specifically, H3K9 methyltransferases bind to CHG methylation to methylate H3K9, and CMT3 binds H3K9me to methylate CHG sites (Du et al., 2012). SUVH4/KRYPTONITE (KYP), SUVH5, and SUVH6, have been identified as H3K9 methyltransferases important in CHG methylation maintenance (Ebbs and Bender, 2006; Ebbs et al., 2005; Jackson et al., 2002; Malagnac et al., 2002). The triple mutant *suvh4/suvh5/suvh6* showed CHG loss similar to that in the *cmt3* mutant, suggesting that SUVH4/5/6 work together to control CMT3 activity (Ebbs and Bender, 2006). Until recently, it has been thought that RdDM is mainly responsible for CHH maintenance, however, defects in RdDM do not eliminate CHH methylation globally, suggesting that there must be a protein involved in CHH maintenance (Cokus et al., 2008; Lister et al., 2008). CMT2, a homolog of CMT3, was identified as a CHH methylation-maintaining DNA methyltransferase (Zemach et al., 2013). Due to the conserved functional domains of CMT3 and CMT2, CMT2 has been proposed to function similarly to CMT3 (Zemach et al., 2013). Moreover, SUVH4/KYP and SUVH6 bind to CHH methylation as well as CHG (Ebbs and Bender, 2006), suggesting that the maintenance of CMT2-mediated CHH methylation involves a reinforcing loop with known H3K9 methyltransferases in the CMT3 pathway.

Regulation of transposon DNA methylation in open and closed chromatin

Recent studies identified that RdDM maintains non-CG methylation in euchromatic transposons and the repetitive edges of heterochromatic transposons (Zemach et al., 2013). DDM1-mediated DNA methylation occurs predominantly in the bodies of long heterochromatic transposons. The loss of CHH methylation due to the RdDM defects is positively correlated with euchromatic modifications but anticorrelated with heterochromatic modifications such as H3K9me2. The same correlation of DDM1-mediated DNA methylation contrast those of RdDM, suggesting that heterochromatin requires DDM1 for DNA methylation and at the same time inhibits access of RdDM machinery (Zemach et al., 2013). Pol IV and V in RdDM are plant specific RNA polymerases that contain subunits paralogous to mRNA-generating Pol II (Haag et al., 2012; Wierzbicki et al., 2008). Since Pol II is active in euchromatin and its access is inhibited by heterochromatin, it is likely that

the ancestral proteins of Pol IV and Pol V functioned in a similar environment and that RdDM is evolved to function in euchromatic regions (Ream et al., 2009)

Interestingly, sRNAs generated by Pol IV also accumulate in heterochromatic transposons (Zemach et al., 2013). This observation raises a question: Can RdDM access heterochromatin if heterochromatin becomes more accessible? Complete decondensation of chromatin naturally occurs in the vegetative nucleus (Schoft et al., 2009), a companion cell responsible for generating the pollen tube required for directing sperm to an egg for fertilization. In the vegetative cell, H3K9me2 is significantly reduced, and DDM1 is inactive (Schoft et al., 2009; Slotkin et al., 2009). Also, transposons are activated (Slotkin et al., 2009), suggesting that a loss of heterochromatic features induces chromatin decondensation and allows the binding of transcriptional machinery to heterochromatin. Unexpectedly, genome-wide DNA methylation maps identified that the global DNA methylation level is generally retained in the vegetative cell (Calarco et al., 2012; Ibarra et al., 2012). Specifically, heterochromatic transposons show non-CG hypermethylation compared to sperm (Ibarra et al., 2012; Schoft et al., 2009) presumably facilitated by RdDM. Together, these results strongly suggest that chromatin organization specifies RdDM affinity.

1.2.2 DNA demethylation in plants

DNA methylation is considered a highly stable and heritable modification, but the methylation marks can be removed by active and passive DNA demethylation mechanisms. In plants, DNA demethylation plays a key role in gene activation, gene imprinting, decondensation of rDNA chromatin, and abiotic/biotic stress responses (Zhu, 2009). Moreover, a mutation in the DNA demethylation pathway is embryo lethal (Choi et al., 2002), emphasizing that proper regulation of DNA demethylation is essential for development.

Active DNA demethylation is facilitated by DNA glycosylases. In plants, DNA glycosylase DEMETER (DME) family proteins, DME, REPRESSOR OF SILENCING (ROS1), DEMETER-LIKE 2 (DML2), and DML3, remove DNA methylation via the base excision repair (BER) pathway (Gehring et al., 2006; Jacobs and Schär, 2012; Penterman et al., 2007). In the BER pathway, DNA glycosylases excise a methylated base on a double stranded DNA. This produces an abasic apurinic/apyrimidinic site (AP site), and the phosphate backbone of the AP site is nicked by an AP endonuclease to generate a 3' hydroxyl. This nick is used as a priming site for DNA polymerase to bind and to fill the DNA gap with an unmethylated cytosine. Mutations in DME, ROS1 and ROS1/DML2/DML3 (RDD) result in locus-specific hypermethylation instead of global hypermethylation (Hsieh et al., 2009; Penterman et al., 2007), suggesting that DME family proteins have specific targets.

Moreover, INCREASE IN DNA METHYLATION 1 (IDM1), an acetyltransferase, is identified as a key component of active demethylation (Qian et al., 2012). IDM1 is predicted to have a methylation binding domain and an N-acetelytransferase domain for recognizing methylated DNA and acetylating H3K18 and H3K24, respectively (Qian et al., 2012). Histone acetylation induces a more relaxed chromatin configuration in eukaryotic species (Görisch et al., 2005; Kodama et al., 2007). Therefore, IDM1 likely creates a more accessible chromatin structure for the binding of a DNA glycosylase and its partner proteins. Loss of IDM1 results in hypermethylation of DME family protein targets, and removal of both IDM1 and ROS1 resembles the *ros1* mutant, indicating that these proteins are in the same pathway (Qian et al., 2012). Moreover, an *idm1* mutation leads to moderate hypermethylation instead of the full hypermethylation shown in *ros1* (Qian et al., 2012), suggesting that another chromatin remodeler is required to create a more permissive chromatin configuration for DME family proteins to bind to their target loci.

1.3 Chromatin architecture matters

Despite recent advances in identifying key chromatin components, there are many remaining questions on how chromatin components together regulate transcription within an organism and how this regulation varies across species.

My dissertation initiated with the following question: how is DNA methylation regulated in plants? To address this question, I examined genome-wide chromatin structure using microarray and whole-genome sequencing. Specifically, I investigated DNA demethylation and chromatin-binding protein mediated DNA methylation patterning in two model plants: rice and *Arabidopsis*.

The second chapter of this work describes DNA demethylation in the rice seed. This work identifies the DNA methylation map of rice seed tissues (endosperm and embryo). The third chapter of this work delves into the relationship between H1 and chromatin components including DNA methylation and nucleosomes in *Arabidopsis*. This study unravels the conundrum of previous studies conducted on H1 and DNA methylation. Finally, the fourth chapter describes a new relationship between H1 and DDM1 in mediating DNA methylation.

Chapter 2:

Local DNA hypomethylation activates genes in rice endosperm

The following chapter has been published as a peer reviewed article in *Proceedings of the National Academy of Sciences of the United States of America*.

Zemach A*, Kim MY*, Silva P, Rodrigues JA, Dotson B., Brooks MD, and Zilberman D (2010) "Local DNA hypomethylation activates genes in rice endosperm." *PNAS*, 107:18729-18734. *These authors contributed equally.

2.1 Abstract

Cytosine methylation silences transposable elements in plants, vertebrates, and fungi but also regulates gene expression. Plant methylation is catalyzed by three families of enzymes, each with a preferred sequence context: CG, CHG (H = A, C, or T), and CHH, with CHH methylation targeted by the RNAi pathway. *Arabidopsis* endosperm, a placenta-like tissue that nourishes the embryo, is globally hypomethylated in the CG context while retaining high non-CG methylation. Global methylation dynamics in seeds of cereal crops that provide the bulk of human nutrition remain unknown. Here, we show that rice endosperm DNA is hypomethylated in all sequence contexts. Non-CG methylation is reduced evenly across the genome, whereas CG hypomethylation is localized. CHH methylation of small transposable elements is increased in embryos, suggesting that endosperm demethylation enhances transposon silencing. Genes preferentially expressed in endosperm, including those coding for major storage proteins and starch synthesizing enzymes, are frequently hypomethylated in endosperm, indicating that DNA methylation is a crucial regulator of rice endosperm biogenesis. Our data show that genome-wide reshaping of seed DNA methylation is conserved among angiosperms and has a profound effect on gene expression in cereal crops.

2.2 Introduction

Roughly 150 million years ago, flowering plants diverged to form the two dominant extant lineages, monocots and dicots (Hedges et al., 2006). *Arabidopsis*, the preeminent plant genetic system, is a dicot, whereas cereal crops, such as rice, wheat, and maize, that feed much of the world are monocots. In both plant groups, pollen grains contain two sperm nuclei, one of which fertilizes a diploid central cell to give rise to triploid endosperm (Huh et al., 2008). *Arabidopsis* endosperm is consumed by the developing embryo, whereas cereal endosperm persists and makes up the bulk of the mature seed— a developmental difference of particular practical importance (Chaudhury et al., 2001). Developing seeds are genetic battlegrounds on multiple fronts: parents are proposed to be in conflict over resource allocation (Huh et al., 2008), whereas the embryo must repress parasitic transposable elements to prevent damage to the genome. *Arabidopsis* endosperm DNA methylation participates in both conflicts (Gehring et al., 2009; Hsieh et al., 2009; Huh et al., 2008; Law and Jacobsen, 2010).

DEMETER (DME) is expressed in the *Arabidopsis* central cell before fertilization, leading to extensive hypomethylation of the maternal genome (Gehring et al., 2009; Hsieh et al., 2009; Huh et al., 2008). This methylation difference

between the maternal and paternal genomes in endosperm—a form of genomic imprinting—causes differential expression of a number of genes, depending on parent of origin (Gehring et al., 2009; Huh et al., 2008). Imprinted expression is generally explained in terms of conflict between parents over resource allocation, with male-expressed genes maximizing resource extraction from the female and female- expressed genes counteracting this drive (Huh et al., 2008).

Most of our knowledge about DNA methylation in plant seeds is derived from *Arabidopsis*. Processes involving genetic conflict tend to evolve rapidly (Swanson and Vacquier, 2002), and therefore, methylation dynamics in cereal seeds may be quite different. Here, we use deep bisulfite sequencing to examine DNA methylation in rice seeds. Wild-type rice endosperm methylation patterns—globally reduced non-CG methylation and local CG hypomethylation—resemble those of DME-deficient *Arabidopsis* endosperm, a finding consistent with lack of DME in monocots. Reduced endosperm methylation is common in genes with preferential endosperm expression, indicating that demethylation is a major mechanism for gene activation in rice endosperm. Short transposons are hypermethylated at CHH sites in embryo, suggesting that endosperm demethylation functions to immunize the embryo against transposons through small RNAs.

2.3 Results

2.3.1 Global genomic methylation patterns are similar across rice tissues

To learn how DNA methylation regulates cereal seed genomes, we quantified methylation in rice embryos, endosperm and seedling shoots and roots first by sequencing bisulfite-converted genomic DNA (bisulfite treatment converts unmethylated cytosine to uracil) to 11- to 15-fold coverage of the nuclear genome (Table 2.1). The aggregate methylation patterns in all tissues are very similar to those of mature rice leaves (Zemach et al., 2010) as well as those of *Arabidopsis* (Law and Jacobsen, 2010)—CG methylation is common in gene bodies, except near the transcription start and termination sites, whereas transposons are methylated in all sequence contexts (Figure 2.1). Overall CG methylation patterns and levels are indistinguishable between embryos, shoots, roots, and leaves (Figure 2.1A and B). CHG methylation increases modestly with age of the tissue: lowest in embryos, higher in young shoots and roots, and highest in mature leaves (Figure 2.1C and D), consistent with reports of increased methylation in older tissues of maize and petunia (Martienssen et al., 1990; Meyer et al., 1992). CHH methylation is also higher in leaves than in seedling tissues (Figure 2.1E and F).

2.3.2 Short transposons are hypermethylated at CHH sites in rice embryo

Average CHH methylation of embryo transposons is higher than in seedlings near the points of alignment but indistinguishable past 1 kb into the element (Figure 2.1F), a pattern caused by differential methylation of short and long transposons (Figure 2.2A and B). Transposons longer than 1kb show the same methylation levels in embryos and seedlings, whereas shorter elements such as miniature inverted-repeat transposable elements (MITES) and short interspersed nuclear elements (SINES) are hypermethylated in embryos (Figure 2.2A and B). The abundance of CHH methylation in short transposons led us to examine whether their genomic distribution accounts for the spike in CHH methylation upstream of genes (Figure 2.1E). MITEs, the most abundant short elements in rice (some of which are active) (Jiang et al., 2003), preferentially occur near genes (Bureau and Wessler, 1994). MITE distribution indeed closely parallels that of CHH methylation (Figure 2.2C). MITE frequency at 5' and 3' ends of genes is directly correlated with gene transcription, whereas MITE frequency within genes is inversely correlated with transcription (Figure 2.2D). CHH methylation shows a similar distribution, a pattern quite different from CG or CHG methylation (Figure 2.3). Thus the distribution of CHH methylation closely follows that of MITEs.

2.3.3 Global hypomethylation of non-CG sites and local hypomethylation of CG sites in rice endosperm

DNA methylation in rice endosperm is lower in all sequence contexts compared with all other tissues that we examined: CG methylation is about 93% of that of embryos, whereas CHG and CHH methylation is lower by about two-fold and five-fold, respectively (Figures 2.1, 2.2A, and Table 2.1). The decrease in CG methylation affects gene bodies, gene-adjacent regions, and transposons (Figure 2.1), which may reflect *Arabidopsis*-like even hypomethylation of the entire genome (Hsieh et al., 2009) or might be because of severe demethylation of specific loci. To identify differentially methylated rice sequences, we calculated fractional methylation in each context within 50 bp windows, subtracted one dataset from another in all pair-wise combinations, and identified loci with significant methylation differences between tissues (Figure 2.4, 2.5 and Table 2.1). CG methylation of most loci is unchanged in rice endosperm (Figure 2.4A), with hypomethylation restricted to specific domains (Figure 2.5), whereas non-CG methylation is reduced throughout the genome (Figures 2.4B, C, and 2.5). Gene bodies, transcriptional start site

proximal regions, and MITEs show decreases in endosperm CG methylation, whereas long transposons exhibit virtually no CG hypomethylation (Figure 2.4D and E). Long transposons also lose less CHG methylation in endosperm than other sequences (Figure 2.4D and E). There are modest (193-1,799 loci) CG methylation differences between tissues other than endosperm, with leaves most similar to seedling shoots and seedling shoots most similar to seedling roots (Table 2.2), suggesting that age and tissue type influence methylation patterns. CG methylation differences are much greater when endosperm is considered with 25,655-29,969 loci hypomethylated in endosperm (Table 2.2).

2.3.4 DEMETER orthology group is restricted to dicots

Our data show that a major reduction of DNA methylation occurs in the endosperm of rice, but the specifics of demethylation are quite different compared with *Arabidopsis* (Gehring et al., 2009; Hsieh et al., 2009), leaving open the question of whether global demethylation has a common evolutionary origin among angiosperms. A plausible answer is suggested by the fact that the methylation landscape of rice endosperm closely resembles that of *Arabidopsis* endosperm with a mutation in the DEMETER (DME) DNA glycosylase (Hsieh et al., 2009). DME is required for global CG demethylation in *Arabidopsis* endosperm (Gehring et al., 2009; Hsieh et al., 2009), and its loss of function leads to increased CG methylation and decreased non CG methylation (Hsieh et al., 2009) very similar to that of wild-type rice (Figure 2.6). DME and its three *Arabidopsis* homologs that function primarily outside the seed (Penterman et al., 2007; Zhu, 2009) belong to three distinct orthology groups within flowering plants: *DME*, *ROS1*, and *DML3* (Figure 2.7). The *DME* orthology group extends to monkey flower (*Mimulus guttatus*) (Figure 2.7), a basal dicot that diverged from *Arabidopsis* about 125 million years ago (Hedges et al., 2006), suggesting that DME function may be conserved across dicots. However, monocots like rice lack *DME* orthologs (Figure 2.7), and therefore rice endosperm hypomethylation must rely on *ROS1* and/or *DML3* orthologs or alternate biochemical mechanisms. The similarity between wild-type rice and DME-deficient *Arabidopsis* suggests that the overall process of extensive endosperm demethylation is likely conserved between monocots and dicots, with the observed differences caused by divergent evolution of the *DME/ROS1/DML3* family.

2.3.5 CG and CHG hypomethylation is a major mechanism of gene activation in rice endosperm

Gene expression in *Arabidopsis* endosperm is significantly linked to DNA methylation—those genes with reduced DNA methylation upstream of the transcriptional start site tend to be more expressed in endosperm (Hsieh et al., 2009). The association is, however, not very strong with most endosperm preferentially expressed genes not directly activated by removal of DNA methylation (Gehring et al., 2009; Hsieh et al., 2009). To examine the situation in rice, we measured gene expression in the same tissues that we used to quantify DNA methylation (embryos, endosperm, and seedling shoots, and roots) using tiling microarrays. We identified a stringent group of 165 genes with a strong preference for endosperm expression by selecting only those genes with fourfold or greater RNA levels in endosperm compared with each of the other three tissues as previously described (Xue et al., 2009). We similarly identified a control group of 153 genes preferentially expressed in embryo.

Genes preferentially expressed in embryo have similar methylation patterns in all tissues (Figure 2.8). In contrast, genes preferentially expressed in endosperm are hypermethylated in embryos, shoots, roots, and leaves on average (Figure 2.8E, F and G). Also, endosperm-preferred genes showed CG and CHG methylation loss in endosperm (Figure 2.8D and E), with 42% (69 genes) displaying a significant decrease (p-value $<10^{-7}$, Fisher's exact test) in CG methylation within 100 bp of the transcription start site or in CHG methylation within the gene body, and 23% (38 genes) exhibiting both. Only 6% (9 genes) of embryo-preferred genes showed a significant decrease in CG or CHG methylation in endosperm, and none showed both even at a highly significant difference (p-value $<10^{-11}$,Fisher's exact test).

For a global comparison of methylation and transcription changes, we aligned all genes according to our microarray data from those most expressed in endosperm versus embryo to those lease expressed in endosperm versus embryo. Then, we displayed embryo methylation levels and the difference between embryo methylation levels embryo and endosperm levels as heat maps (Figure 2.9). Endosperm-expressed genes have higher levels of embryo CG methylation near the transcriptional start sites (Figure 2.9A) and higher embryo CHG methylation throughout the gene body (Figure 2.9B), with reduced level of methylation in endosperm. CHH methylation shows a weak, if any, correlation with embryo and endosperm expression differences (Figure 2.9C). Our data suggest that DNA hypomethylation is a major mechanism for activation of genes in rice endosperm. Endosperm-preferred genes that are apparently activated by DNA demethylation include precursors of glutelin, which accounts for about 70% of rice endosperm protein (Qu et al., 2008), and carbohydrate enzymes that synthesize endosperm

starches (Figures 2.5), genes that create the nutritive molecules relied on by germinating seedlings and much of the human population.

2.4 Discussion

2.4.1 DNA demethylation in rice seeds

It was previously suggested that DNA hypomethylation in *Arabidopsis* endosperm is a mechanism to enhance transposon silencing in the embryo (Hsieh et al., 2009), a hypothesis supported by an abundance of transposon-derived small RNAs in the endosperm (Mosher et al., 2009). Our rice data are fully consistent with this hypothesis. The greatest amount of CHH methylation is found in short rice transposons, consistent with the observation that methylation of short transposons is more dependent on the RNAi system (Tran et al., 2005). Short transposons lose the most CHH methylation in rice endosperm (Figure 2.1F) and are the only elements hypermethylated in embryo (Figure 2.2A), suggesting that demethylation and activation of transposons in endosperm lead to hypermethylation and silencing in the embryo through the RNAi pathway. Considering that the major targets of rice CHH methylation are MITEs, which tend to be found near the start sites of active genes, this system is likely of particular importance for proper functionality of the rice genome.

It is strongly suggest that the DNA demethylation mechanism in seed development is conserved in both monocots and dicots. Although a rice DME homolog was not found in our study, the rice ROS1 homolog, ROS1a, is identified to catalyze DNA demethylation in rice seeds (Ono et al., 2012). ROS1a is expressed in the central cell and likely responsible for the loss of DNA methylation in endosperm (Ono et al., 2012). Also, the seeds with a maternal *ros1a* mutant allele and the wild-type paternal *ROS1a* allele exhibited similar endosperm defects to that of *Arabidopsis* (Choi et al., 2002; Ono et al., 2012), indicating that ROS1a appears to function as DME in rice female gametophytes.

Moreover, we find that major resource genes are activated by hypomethylation in endosperm and ROS1a in central cell presumably mediates the hypomethylation (Ono et al., 2012). However, if the expression of major resource genes is confined to the maternal genome, our observations would seem inconsistent with the prevalent parental conflict theory, which predicts that genes that increase resource allocation should be expressed from the paternal genome. A recently published study investigated paternal and maternal DNA methylation in endosperm by analyzing a cross of two different rice cultivars, Kitaake and Nipponbare (Rodrigues et al.,

2013). This study identified that paternally expressed genes are also maternally demethylated (Rodrigues et al., 2013). Unlike maternally expressed genes that lose DNA methylation at the transcription start site, the paternally expressed genes lose DNA methylation in the body of genes (Rodrigues et al., 2013), suggesting that DNA methylation in gene body and promoter have different functions.

2.4.2 Establishing genomic imprinting before and after fertilization

A number of monocot imprinted genes apparently activated by selective maternal demethylation have been identified (Gutiérrez-Marcos et al., 2006; Haun et al., 2007; Jahnke and Scholten, 2009). One study in maize identified that *FIE2* shows a monoallelic maternal expression in early endosperm development but not in the gametophyte (Danilevskaya et al., 2003; Gutiérrez-Marcos et al., 2006). This study implies that gene imprinting also occurs during or after fertilization and methylation patterns in the endosperm do not simply reflect DNA methylation patterns present in the gametes. Whether and to what extent the demethylation that we observe occurred before or after fertilization are urgent issues for further study.

It is very difficult to investigate DNA methylation and gene expression in *Arabidopsis* gametes due to technical difficulties in isolating gametes. The female gametes (the egg and central cell) are especially difficult to purify due to being buried inside a complex structure of the ovule which exacerbates the problem of isolating cells without contamination. The nuclei of male gametes (the sperm and vegetative cell) have been isolated via Fluorescence Activation Cell Sorting (FACS) (Calarco et al., 2012; Ibarra et al., 2012), but the mechanized cell sorting system still faces a major problem of cross contamination. Techniques to isolate gametes manually have been developed and optimized in many crop plants such as maize and rice (Khalequzzaman and Haq, 2005; Kranz et al., 1991; Uchiumi et al., 2006). Manual isolation secures the isolation of the highest purity of cells when compared with other techniques such as FACS. Therefore, rice is a great candidate for further studying gamete DNA methylation and investigating gene imprinting before fertilization.

2.4.3 Implications for engineering transgenic cereal crops

Our data have important implications for engineering of transgenic cereal crops. Glutelin promoters have been investigated for their suitability to drive transgene expression in rice endosperm, and strong endosperm-specific promoters are

generally highly sought (Qu et al., 2008). However, our data suggest that a substantial fraction of endosperm-specific expression is caused by hypomethylation, and transgenic constructs may not recapitulate endogenous expression patterns. Understanding the epigenetic mechanisms regulating such promoters will allow for a more informed design of transgenic lines.

Considering the difficulty of targeted mutagenesis in plants (Li et al., 2007b), RNAi is an attractive mechanism for silencing unwanted endogenous genes. However, the extremely low levels of CHH methylation in rice endosperm indicate that the functionality of the RNAi system is altered in this tissue. Small RNAs may be transported out of the endosperm to immunize the embryo, or the link between RNAi and DNA methylation may be weakened. In either case, targeting RNAi to promoters for the purpose of transcriptional repression may be ineffective. Similarly, transgenes inactivated by DNA methylation in other tissues may be reactivated in endosperm.

2.5 Materials and methods

Rice material

Endosperm and embryo were harvested from rice (*Oryza sativa* ssp. *japonica* cultivar *Nipponbare*) plants grown in greenhouse. Rice endosperm and embryo from 20-50 seeds at milky stage were collected. The embryo was washed in 1% Triton X-100 to dissolve eliminate tissue contamination. The samples were harvested and frozen in liquid nitrogen and stored in -80ºC until use. Rice seedlings were grown in liquid culture. The rice seed coat of 10-15 seeds was removed then sterilized in 70% ethanol for 5 minutes and 20% bleach for 15 minutes on a shaker. The seeds were washed thoroughly with water then grown in sterile Gamborg's B-5 medium (Caisson labs, adjusted to pH 5.0) supplemented with 5uM biotin (80mL culture in 250 mL flask). The seedlings were grown under a continuous light for 4 weeks under a gentle shake at approximately 80 rpm.

Methylated DNA immunoprecipitation (MeDIP) – chip

Genomic DNA was extracted from 50mg of fresh or frozen roots were ground then resuspended in 2x CTAB buffer containing 2% CTAB (hexadecyltrimethylammonium bromide), 100mM Tris-HCl pH=8, 20mM EDTA, 1.4M NaCl. The suspension was incubated at 65ºC for 1-3 hours then centrifuged at maximum speed for 10 minutes. The supernatant was transferred to a new tube and purified by phenol::chloroform extraction. The aqueous phase was transferred then 0.7 volumes of isopropanol. The pellet was resuspended in TE containing

100mM Tris-HCl pH=8 and 1mM EDTA then treated with RNAse A at 37ºC for 1 hour to remove RNA. DNA was purified by phenol::chloroform extraction and precipitated by 3 volumes of 100% ethanol and 1/100 volume of 3M sodium acetate.

4ug of sheared DNA to 200-1000 bp in 450ul TE was denatured for 10 minutes in boiling water then immediately cooled for 10 minutes on ice. IP buffer containing 10mM sodium phosphate pH 7.0, 140mM NaCl, and 0.05% Triton X-100 and 10ul of 5mC antibody was added and incubated at 4ºC for 2 hours on a rotator. 40ul of Dynabeads (Invitrogen) was added to the reaction and incubated at 4ºC for 2 hours on a rotator. The beads were washed with IP buffer twice. The beads were resuspended in 250ul IP buffer with 7ul proteinase K and incubated at 50ºC for 3 hours on a thermoshaker. The DNA was extracted using phenol-chloroform extraction and precipitated with 3 volumes and 1/100 volume of 3M sodium acetate. The enrichment of MeDIP was measured by PCR by amplifying ACTIN as unmethylated control and Ta5 as a methylated control.

The IP and IC was amplified by the GenomePlex Whole Genome Amplification Kit (Sigma) as 1ug of DNA is required for the microarray library construction. For each tissue, two independent replicate libraries were constructed according to the manufacturer's instruction (NimbleGen) and hybridization and data extraction were performed at the Fred Hutchinson Cancer Research Center DNA array facility.

RNA-chip

RNA was isolated using RNeasy Plant Mini Kit (Qiagen). Due to high starch content, we used Buffer RLC that contains guanidine hydrochloride for greater cell disrubtion during lysis step. The genomic DNA was removed on the column using RNase-Free DNase Set (Qiagen). cDNA was generated using SuperScript III Reverse Transcriptase Kit (Invitrogen) and labeled as described above in MeDIP-chip. Two biological replicates for each tissue were labeled with Cy5 and cohybridized with sonicated genomic DNA labeled with Cy3. The two replicates were averaged, and outlier probes were removed by median smoothing (three-probe window). An expression score for each gene was calculated by averaging the signal of all probes within the gene's exons.

Bisulfite sequencing

About 500 ng of genomic DNA was fragmented by sonication to 200-1kb in size. The fragmented DNA was end repaired and ligated to custom synthesized methylated adapters (Eurofins MWG Operson) according to the manufacturer's instructions for gDNA library construction (Illumina).

The exact sequences of adapter oligos are

5'-ACACTCTTTCCCTACACGACGCTCTTCCGATCT-3' and
3'-GAGCCGTAAGGACGACTTGGCGAGAAGGCTAGp-5'.

Adaptor-ligated libraries were subjected to two treatments of sodium bisulfite conversion using the EpiTect Bisulfite kit (Qiagen) as outlined in the manufacturer's instructions. The bisulfite-converted libraries were then amplified by 18 cycles of PCR using PfuTurboC_xDNA polymerase (Stratagene), the enzyme tolerates uracil in the template. PCR reactions were carried out as follows: 95ºC for 30 seconds, 12-14 cycles of 95ºC for 30 seconds, and 72ºC for 60 seconds. The enriched libraries were either gel purified (~300 bp band) and 10 fragments were cloned for validation. The Sanger sequencing of cloned fragments confirmed whether clones were the full bisulfite converted and had correct adapter sequences. Sequencing on the Illumina platform was performed at the Vincent J. Coates Genomic Sequencing Laboratory at the University of California, USA (UC Berkeley).

2.6 Figures and tables

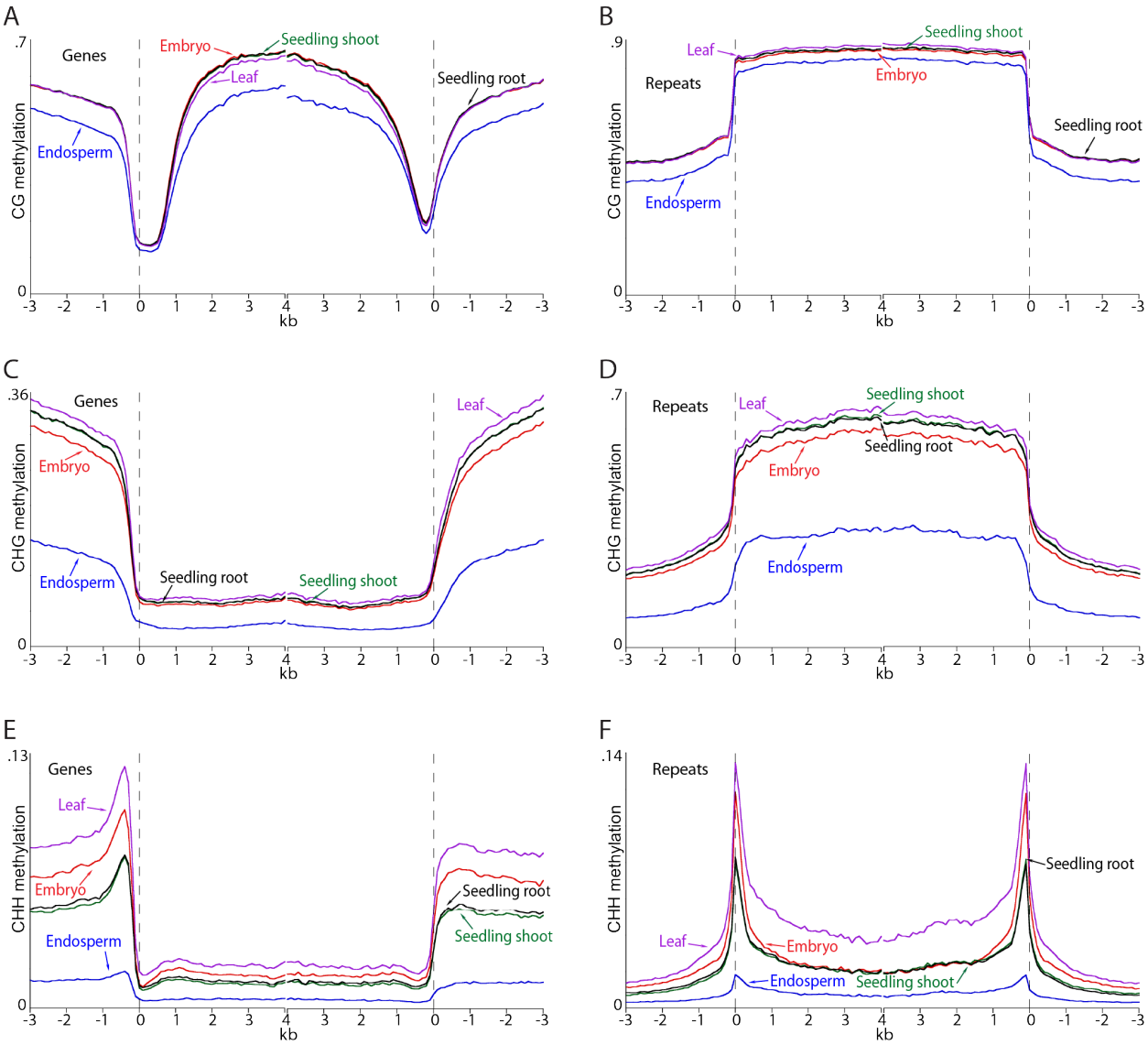

Figure 2.1. Patterns of DNA methylation in rice tissues. Rice genes (A, C, and E) or transposons (B, D, and F) were aligned at the 5' end (left) or the 3' end (right), and average methylation levels of each 100 bp interval are plotted. The dashed line represents the point of alignment.

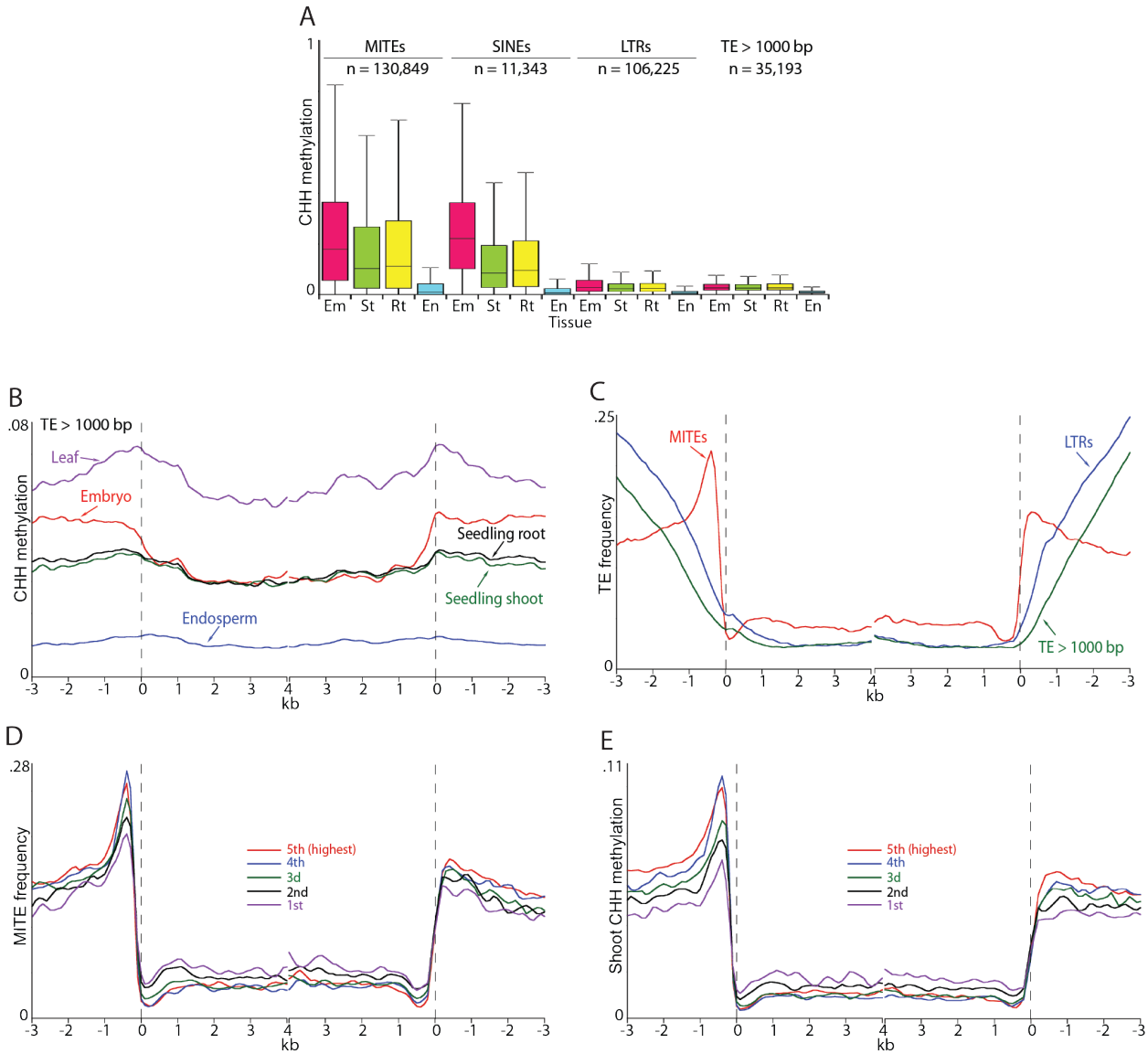

Figure 2.2. MITEs are the predominant targets of CHH methylation. (A) Box plots showing methylation levels of different transposon classes in rice embryo (Em), shoot (St), root (Rt), and endosperm (En). Each box encloses the middle 50% of the distribution, with the horizontal line marking the median. The lines extending from each box mark the minimum and maximum values that fall within 1.5 times the height of the box. MITE mean length = 189 bp, and maximum length = 500 bp. SINE mean length = 141 bp and maximum length = 487 bp. LTR mean length = 855 bp and maximum length = 11 kb. (B) Distribution of CHH methylation in transposons > 1000 bp. Rice transposons longer than 1000 bp were aligned at the 5' end (left side) or the 3' end (right side), and average methylation levels for each 100 bp interval are plotted. The dashed line represents the point of alignment. (C-D) Rice genes were aligned as in Figure 2.1, and transposon frequency (B and C) or average methylation levels (D) for each 100 bp interval are plotted. In C and D, genes were grouped into quintiles by transcription.

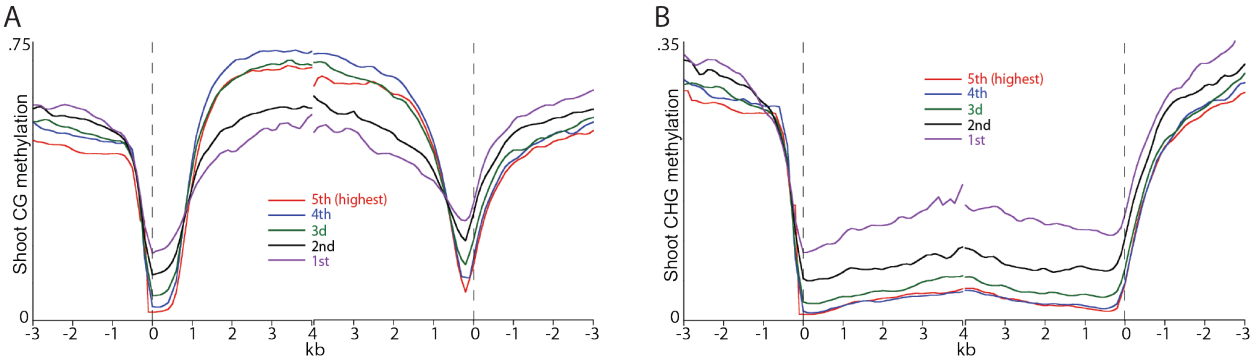

Figure 2. 3.Variation in CG and CHG methylation patterns associated with transcription. (A-B) Rice genes were aligned at the 5' end (left side) or the 3' end (right side), and average CG (A) and CHG (B) methylation levels for each 100 bp interval are plotted. Genes were grouped into quintiles by transcription. The dashed line represents the point of alignment.

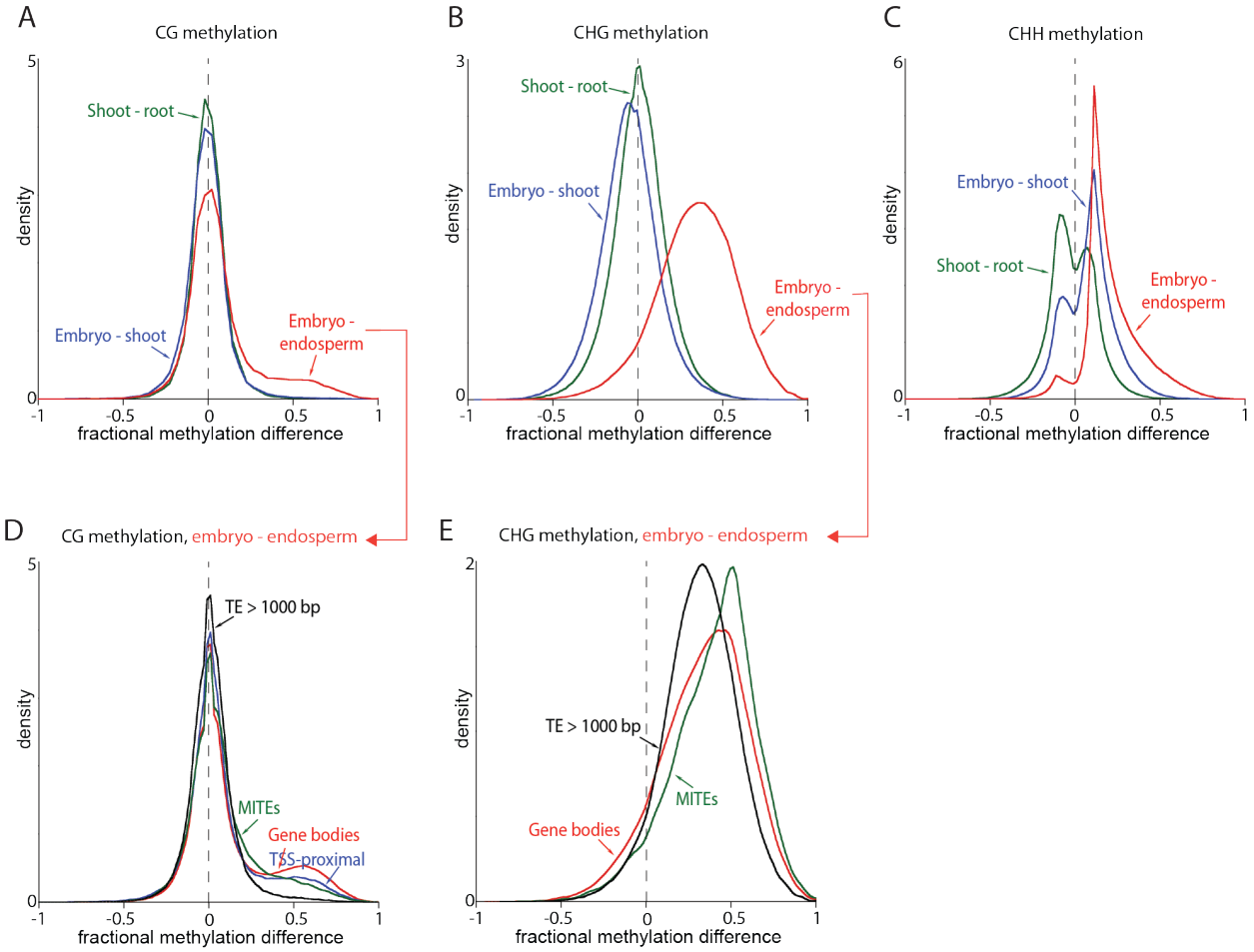

Figure 2.4. Local CG and global non-CG hypomethylation of rice endosperm. Kernel density plots of the differences between embryo and endosperm methylation (red trace), the differences between embryo and seedling shoot methylation (blue trace), and the differences between seedling shoot and root methylation (green trace). Methylation differences for 50 bp windows containing at least 10 informative sequenced cytosines are shown. Differences for windows with fractional CG methylation of at least 0.7, fractional CHG methylation of at least 0.5, and fractional CHH methylation of at least 0.1 in one of the tissues are shown, respectively. (D-E) The 50 bp windows from the embryo-endosperm from (A) and (B) were clustered by their locations in gene bodies (red trace), transcriptional start site (TSS)-proximal sequences (+200 to -500 bp; blue trace), miniature inverted-repeat transposable elements (MITEs, green trace, and transposons longer than 1000 bp (black trace).

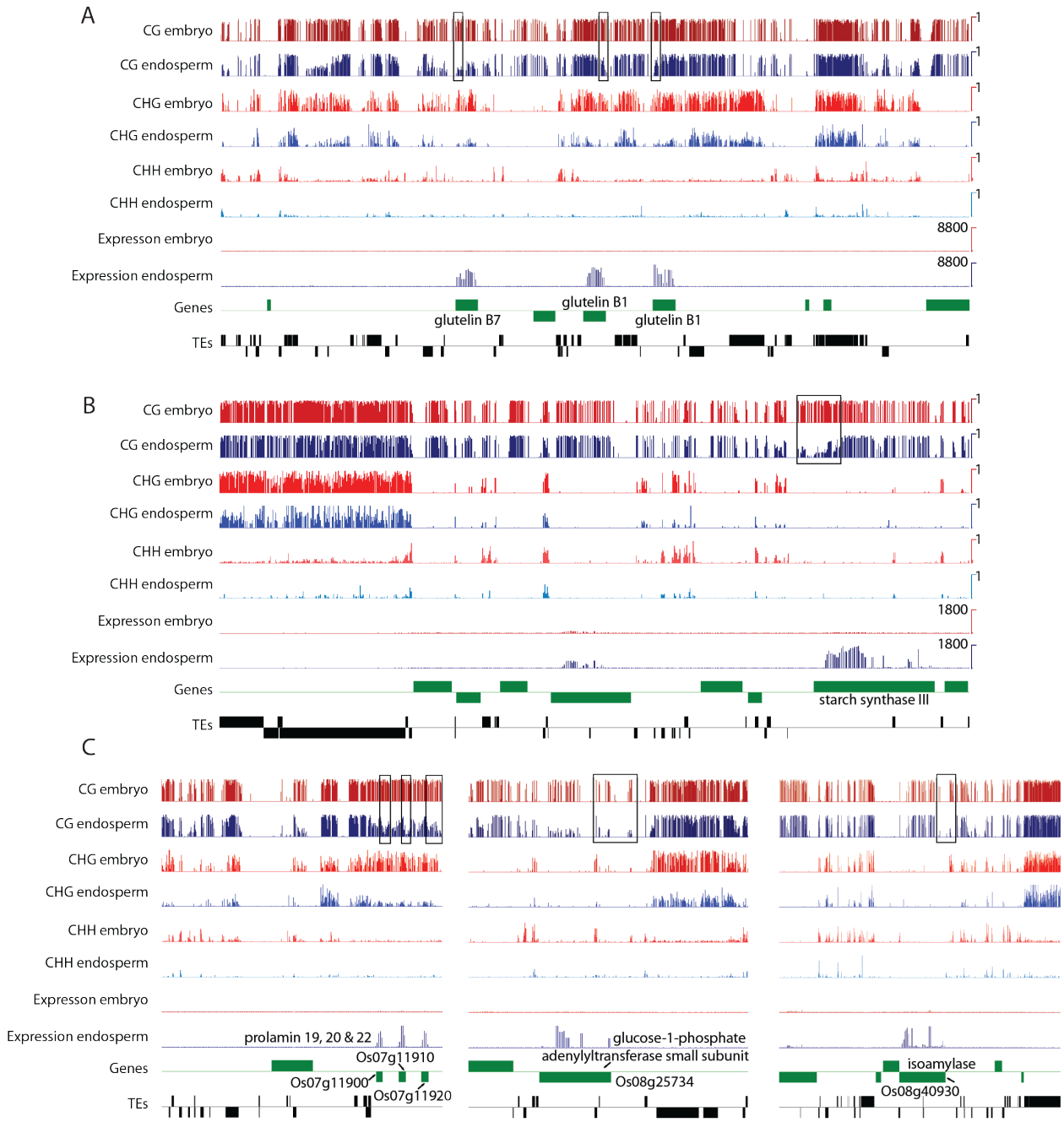

Figure 2.5. Hypomethylation of endosperm-activated rice genes. (A-C) Methylation and transcription of 25 kb regions surrounding the indicated storage protein (A) and starch biosynthesis genes (B-C) in embryo and endosperm. Genes and transposons oriented 5' to 3' and 3' to 5' are shown above and below the line respectively. Endosperm CG hypomethylation overlapping the transcription start site of the relevant genes is highlighted in boxes. Also note that extensive loss of CHG methylation.

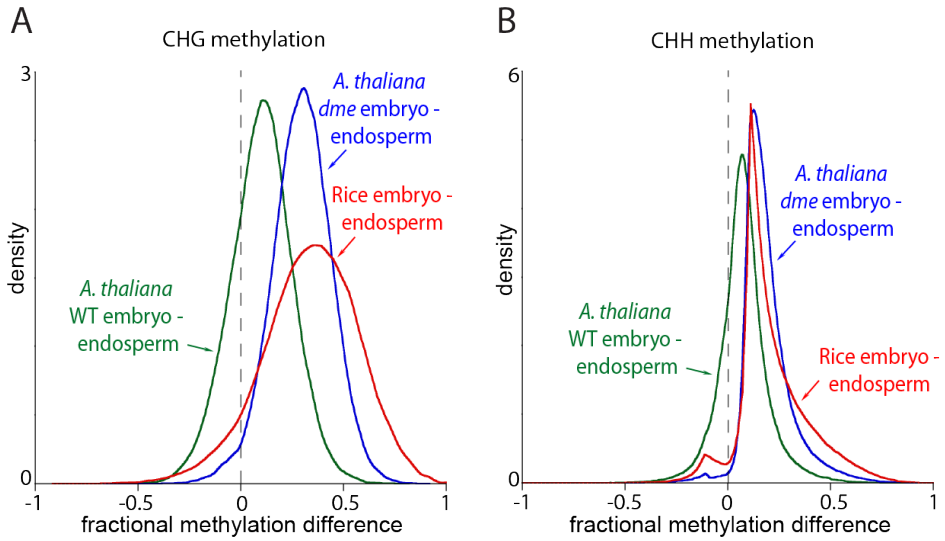

Figure 2.6. Wild-type rice endosperm methylation resembles *Arabidopsis dme* endosperm. Kernel density plots of CHG and CHH methylation differences within 50 bp windows between rice embryo and endosperm (red trace), *Arabidopsis* endosperm from loss-of-function *dme* plants (blue trace), and *Arabidopsis* endosperm from wild-type plants (green trace). Differences for windows with functional CHG methylation of at least 0.5 in one of the tissues are shown in A and differences for windows with fractional CHH methylation of at least 0.1 in one of the tissues are shown in B.

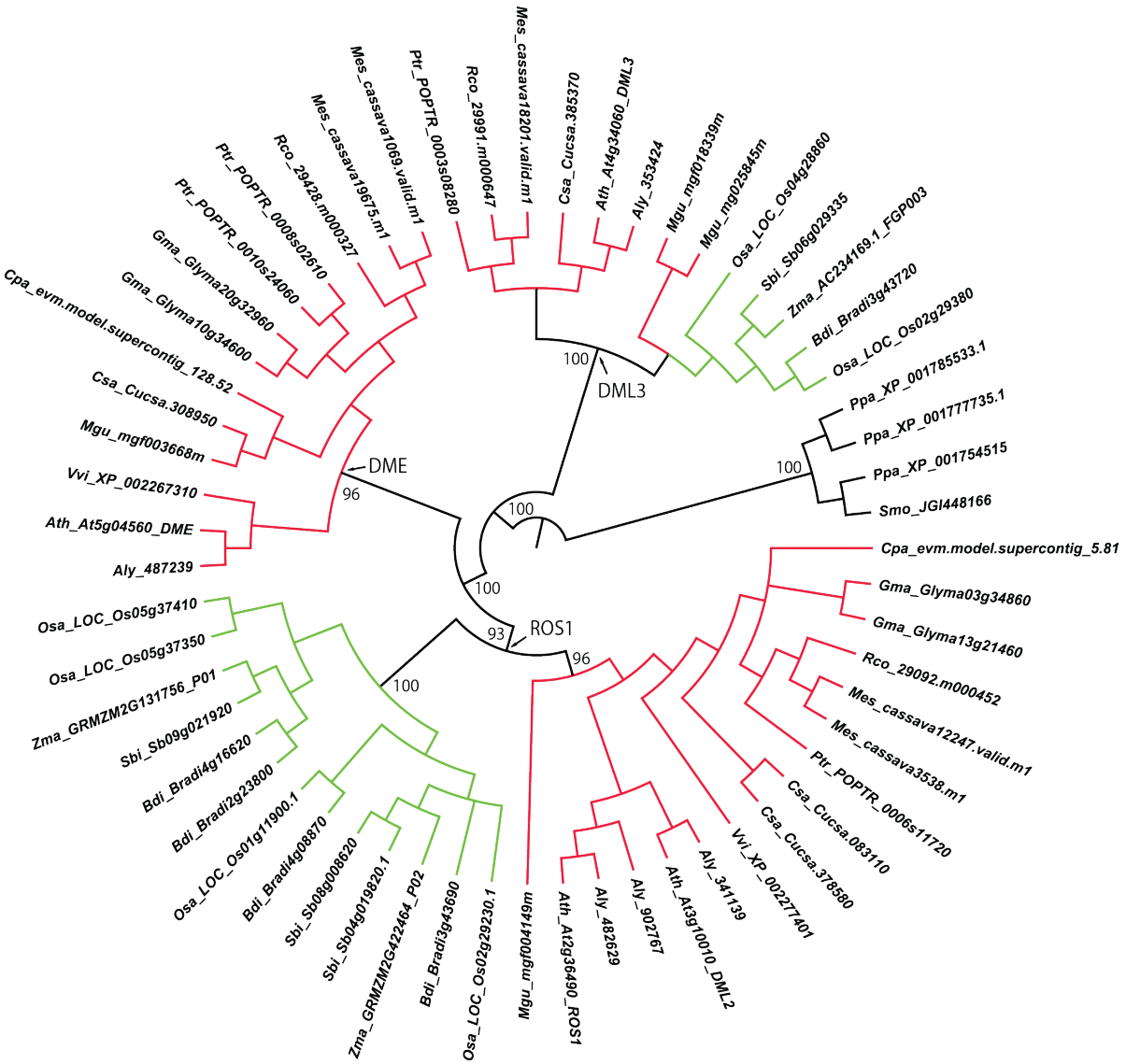

Figure 2.7. Phylogenetic analysis of the DME/ROS1/DML3 glycosylase family. A phylogenetic tree based on conserved domains of glycosylase proteins, with basal land plant (moss and lycophyte) proteins as an outgroup. Posterior probability values (0–100) are shown for key nodes. Monocot, dicot, and moss/lycophyte proteins are colored green, red, and black, respectively. Aly, *A. lyrata*; Ath, *A. thaliana*; Bdi, *Brachpodium distachyon*; Cpa, *Carica papaya* (papaya); Csa, *Cucumis sativus* (cucumber); Gma, *Glycine max* (soybean); Mes, *Manihot esculenta* (cassava); Mgu, *Mimulus guttatus* (monkey flower); Osa, *Oryza sativa* (rice); Ppa, *Physcomitrella patens* (moss); Ptr, *Populus trichocarpa* (poplar); Rco, *Ricinus communis* (castor bean); Smo, *Selaginella moellendorffii* (lycophyte); Sbi, *Sorghum bicolor*; Vvi, *Vitis vinifera* (grape); Zma, *Zea mays* (maize). Protein sequences were downloaded from Phytozome (http://www.phytozome.net), National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov), and Joint Genome Institute (http://www.jgi.doe.gov).

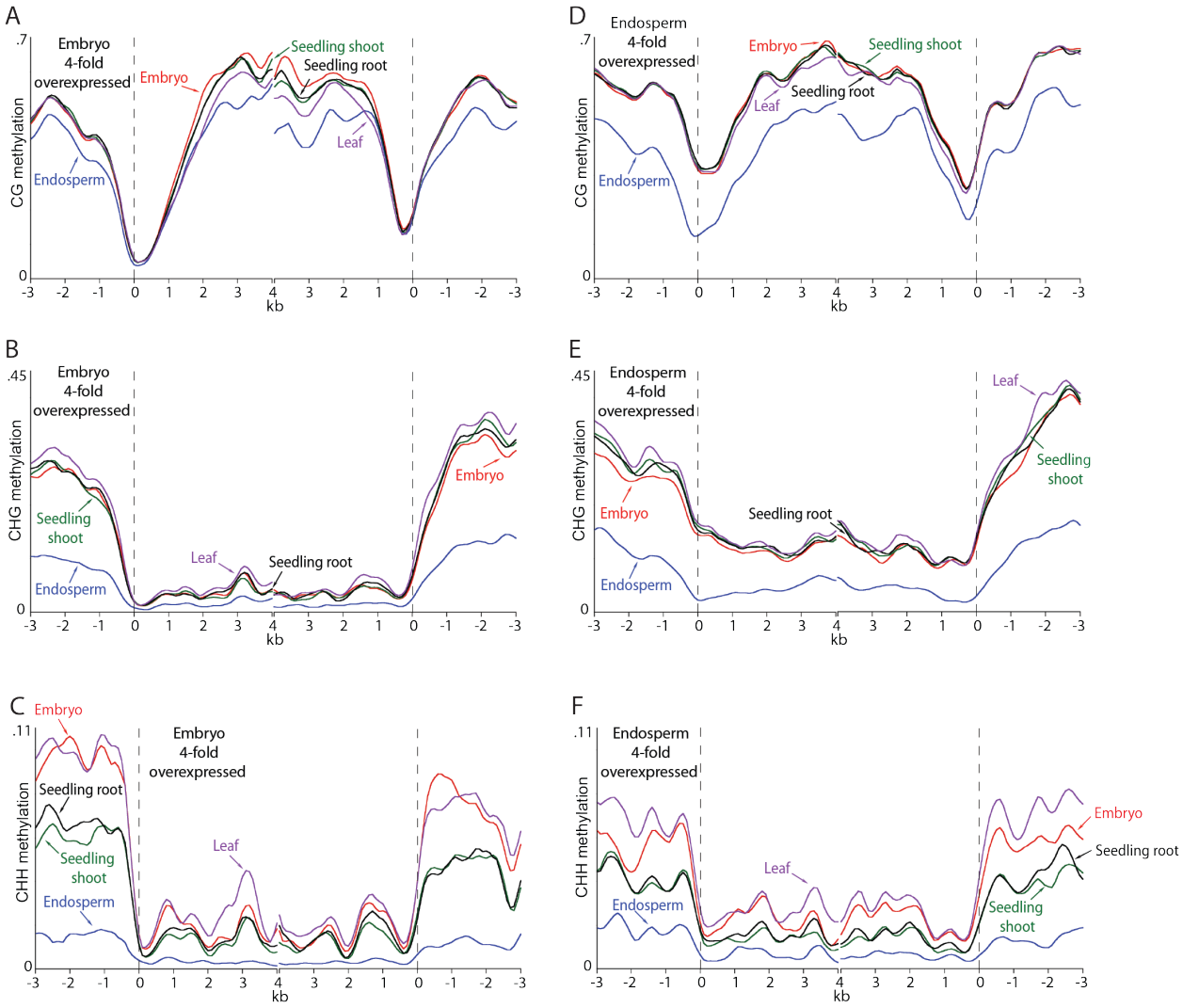

Figure 2.8. Hypomethylation is a major mechanism for activation of rice endosperm genes. (A-F) Embryo-preferred (A-C; n=153) or endosperm preferred (D-F; n=165) genes were aligned as in Figure 2.1, and average methylation levels for each 100 bp interval are plotted.

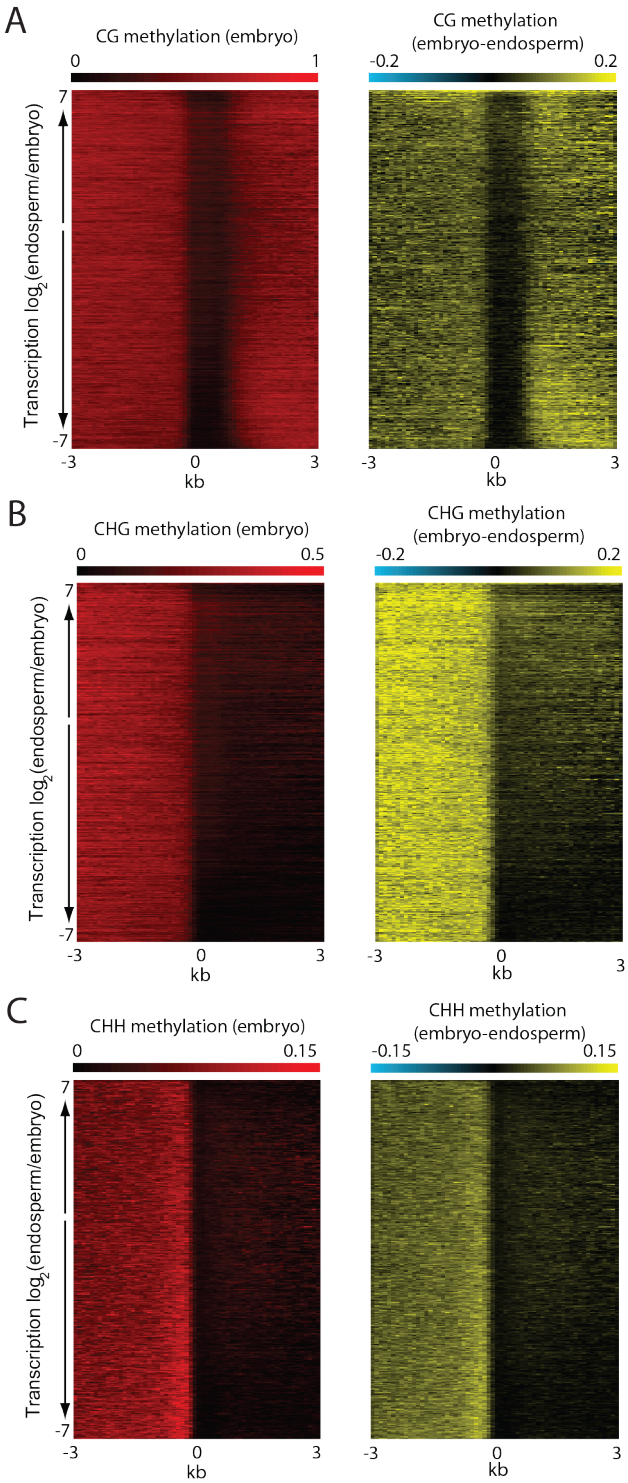

Figure 2.9. Hypomethylation in endosperm versus transcription. All rice genes were aligned at the 5' end and ranked according to the ratio of expression in endosperm over embryo. Embryo methylation is displayed as a heat map in *Left*, and differences between embryo and endosperm are in *Right*.

Tissue	Genome (Mb)	Coverage (X fold)	C (%)	CG (%)	CHG (%)	CHH (%)
Embryo nuclear	372.6	13.5	13.6	38.5	19.9	3.4
Endosperm nuclear	372.6	14.9	9.7	36.0	9.7	0.65
Seedling shoot nuclear	372.6	15.2	12.9	38.8	19.3	2.5
Seedling root nuclear	372.6	11.0	13.0	39.4	19.2	2.7
Embryo chloroplast	0.134	438	0.58	1.66	1.15	0.11
Endosperm chloroplast	0.134	105	1.08	4.36	1.10	0.09
Seedling shoot chloroplast	0.134	403	0.69	2.06	1.48	0.09
Seedling root chloroplast	0.134	90	2.21	6.56	4.60	0.25
Embryo mitochondrion	0.491	628	0.32	0.90	0.52	0.08
Endosperm mitochondrion	0.491	300	0.47	1.78	0.45	0.07
Seedling shoot mitochondrion	0.491	292	0.81	2.51	1.50	0.09
Seedling root mitochondrion	0.491	324	0.64	1.78	1.05	0.16

Table 2.1. Bulk methylation statistics for the nuclear, chloroplast, and mitochondrial genomes.

	Embryo	Endosperm	Seedling shoot	Seedling root	Mature leaf
Embryo	x	631	535	326	897
Endosperm	26,916	x	29,696	28,288	25,655
Seedling shoot	481	562	x	152	193
Seedling root	585	625	389	x	560
Mature leaf	1,799	1,510	792	1,100	x

Table 2.2. Number of hypermethylated loci in rice tissues. Number of loci with CG methylation greater in the tissue listed along the top versus the tissue listed along the left. Locus length ≥ 200 bp; absolute methylation difference ≥ 20%; $p < 10^{9}$ (Fisher's exact test).

Chapter 3:

Histone H1 stabilizes chromatin and inhibits access of DNA methyltransferases

The mutant generation and the DNA methylation mapping sections of following chapter have been published in a peer-reviewed article in *Cell*.

Zemach A*, Kim MY*, Hsieh PH*, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, and Zilberman D (2013) "The nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin." Cell, 153:193-205.
*These authors contributed equally.

3.1 Abstract

Histone H1 binds to linker DNA between nucleosomes and promotes nucleosome compaction, but the biological significance of H1-mediated chromatin changes is poorly understood. Here, we show that H1 preferentially associates with and increases the distance between nucleosomes in transposons. H1 is also present in gene bodies, where its abundance and effect on nucleosomes are anticorrelated with expression. Lack of H1 causes increased DNA methylation of heterochromatic transposons, but leads to reduced methylation of more euchromatic transposons and genes. Our findings reveal that histone H1 is found throughout chromatin with a dual function differentiated by active and inactive chromatin.

3.2 Introduction

Eukaryotic DNA is wrapped around octamers of histones H2.A, H2.B, H3 and H4 to form nucleosomes. Histone H1 binds to the nucleosome core and the linker DNA that separates nucleosomes (Thomas, 1999; Vignali and Workman, 1998) with a possible preference for methylated DNA (McArthur and Thomas, 1996). H1 condenses chromatin and inhibits nucleosome mobility and transcription *in vitro* (Clausell et al., 2009; Laybourn and Kadonaga, 1991; Pennings et al., 1994; Robinson et al., 2006) and is associated with more compact, less accessible and transcriptionally silent chromatin *in vivo* (Ascenzi and Gantt, 1999; Barra et al., 2000; Corona et al., 2007; Fan et al., 2005; Jedrusik and Schulze, 2001; Krishnakumar et al., 2008; Raghuram et al., 2009; Shen et al., 1995). In mammals, H1 is the most divergent class of histone proteins with eleven different subtypes (Happel and Doenecke, 2009). *Arabidopsis* has three known subtypes and two are ubiquitously expressed canonical histone H1 proteins, H1.1 and H1.2. (Gantt and Lenvik, 1991; Wierzbicki and Jerzmanowski, 2005). H1.3 is identified to be stressed inducible and has a distinct structure with a short C-terminus compared to the canonical histone H1 proteins (Ascenzi and Gantt, 1997; 1999).

Loss of H1 has been reported to cause disparate and apparently contradictory changes in genomic DNA methylation. Mice with reduced H1 lose DNA methylation at the regulatory regions of several imprinted genes, but global methylation patterns remain grossly unperturbed (Fan et al., 2005). In contrast, removal of H1 leads to extensive hypermethylation in the fungus *Ascobolus immersus* (Barra et al., 2000). In *Arabidopsis*, loss of H1 has also been reported to cause stochastic methylation gains and losses (Rea et al., 2012; Wierzbicki et al., 2008). Also, plant DNA methylation is mediated by distinct methyltransferase families in three sequence contexts (CG, CHG and CHH, where H is any nucleotide except G) (Law

and Jacobsen, 2010), but which of these methyltransferases is impacted by H1 binding is unknown.

To understand how H1 affects nucleosome organization and DNA methylation *in vivo*, we mapped H1 binding sites throughout the *Arabidopsis* genome and analyzed nucleosome occupancy and DNA methylation in plants with mutated H1 genes. We identified that H1 is enriched in heterochromatin and regulates nucleosome spacing. Also, loss of H1 resulted in hypomethylation in euchromatic elements and hypermethylation in heterochromatic elements, indicating that H1 has a dual function in regulating DNA methylation.

3.3 Results

3.3.1 H1 is enriched in transposons and negatively correlated with gene expression.

Previous cytological studies determined that the heterochromatin was enriched with H1 (Ascenzi and Gantt, 1999; Lever et al., 2000; Misteli et al., 2000). However, it remains unclear whether H1 is also localized in euchromatin due to the low resolution of the assays. To investigate H1 binding sites in more detail and gain insights into H1 function from the genome distribution, we generated a high-resolution map of H1 by sequencing H1-bound DNA fragments.

We tagged H1 using the *in vivo* biotinylation system described previously (de Boer et al., 2003; Furuyama et al., 2006; Mito et al., 2005; Zilberman et al., 2008). In this system, a small 23 amino acids biotin ligase recognition peptide (BLRP) is fused onto the protein of interest. By expressing the transgenic protein with BirA, a bacterial protein biotin ligase (Howard et al., 1985), BLRP will be biotinylated *in vivo*. Then the biotinylated protein of interested can be purified using avidin or streptavidin, the bacterial protein known to have extremely high affinity to biotin. In nature, the avidin-biotin binding is the strongest noncovalent interaction known with the dissociation constant being orders of magnitude lower than any antibody to antigen association (de Boer et al., 2003). For chromatin immunoprecipitation (ChIP), this highly stable binding property of biotin and streptavidin/avidin allowed us to simplify the experiment by eliminating the nuclei extraction step, and thus reduce loss of the sample. Also, extensive washing of ChIP samples significantly lowered nonspecific binding.

We created transgenic lines that express either of the two canonical *Arabidopsis* H1 genes, *H1.1* or *H1.2*, fused to a BLRP and co-expressed the tagged protein with *ACTIN2* driven BirA (Zilberman et al., 2008). Sequencing of H1-associated DNA

revealed that H1.1 and H1.2 have similar genomic distributions (Figure 3.1), suggesting their functional redundancy. The chromosome-level distribution of H1 determined by averaging H1 levels in 200 kb windows showed that H1-associated DNA is most enriched in centromeric regions of each chromosome (Figure 3.1A), consistent with the cytological study in *Arabidopsis* (Ascenzi and Gantt, 1999).

Moreover, H1 is significantly enriched in transposons (Figure 3.1B). The average pattern of H1 enrichment in transposons showed relatively lower H1 levels at the points of alignment (Figure 3.1B). In *Arabidopsis*, > 80% of transposons are shorter than 1 kb (Buisine et al., 2008), and these transposons would only contribute to the patterns of transposon methylation near to the points of alignment. Also, these short transposons have low H3K9me2 (hence euchromatic transposons) compared to longer transposons with high H3K9me2 (heterochromatic transposons) (Zemach et al., 2013). In order to test if the low H1 enrichment at the points of alignment is due to H3K9me2-depleted euchromatic transposons, we clustered transposons based on levels of H3K9me2 (high and low) and generated the average H1 enrichment pattern in transposons. We identified that the euchromatic transposons with low H3K9me2 have very low H1 enrichment, whereas the H3K9me2 enriched heterochromatic transposons are high in H1 levels (Figure 3.2C).

Although the overall enrichment is lower than transposons, genes also have some level of H1 indicating that H1 is not solely a heterochromatic component unlike H3K9me2 (Figure 3.1C). As histone H1 has been associated with transcriptional silencing (Laybourn and Kadonaga, 1991; Shen and Gorovsky, 1996), we further investigated whether enrichment of H1 influences the activity of genes. We clustered genes that are low in H3K9me2 levels into five groups (Figure 3.2D), based on the expression levels, and then determined the average H1 enrichment in each cluster. We identified that the H1 enrichment in gene bodies is anticorrelated with the transcription level—inactive genes have high H1 levels and active genes have low H1 levels (Figure 3.2A and B).

3.3.2 Characterization of the *h1* mutant

To examine the consequences of reduced H1 levels *in vivo*, we crossed *h1.1* and *h1.2* with a T-DNA inserted in first exon and in 5'UTR, respectively (Figure 3.4A). To test if the H1 level is significantly reduced in *h1.1h1.2* (hereafter denoted *h1*), we performed qPCR on *h1* and wild-type using primers that span two exons. We determined that *H1.1* expression level is significantly reduced but *H1.2* expression showed no change compared to wild-type H1.1 and H1.2 expression levels (Figure 3.4B and C). In addition, the stress inducible H1 subtype, H1.3, in wild-type and *h1* mutants is constantly expressed at the basal level, indicating that H1.3 does not compensate for the reduced expression of the canonical H1 proteins (Figure 3.4C).

It is highly likely that the residual *H1.1* mRNA cannot translate functional H1.1 protein due to the T-DNA insertion in the exon. In contrast, H1.2 mRNA from the *h1.2* allele is able to produce a functional protein because the T-DNA insertion is in 5'UTR before the start codon. Due to the possibility that the T-DNA insertion in the *h1.2* allele interferes with translation, we determined the H1.1 and H1.2 protein levels in wild-type, *h1.1*, *h1.2*, and *h1* using anti-H1.1 and anti-H1.2 polyclonal antibodies generated with previously described H1 epitopes (Ascenzi and Gantt, 1997). In *h1* compared to wild-type, H1.1 protein is completely eliminated (Figure 3.4D), consistent with the qPCR result. Only residual H1.2 was detected in *h1*, suggesting that T-DNA in the 5' UTR indeed interfered with translation (Figure 3.4D). Interestingly, H1.2 protein level was significantly lower in *h1* compared to *h1.2* single mutant, making it likely that our anti-H1.2 cross-reacts with H1.1 proteins (Figure 3.4D).

3.3.3 H1 increases the distance between nucleosomes in *Arabidopsis*

Adjacent nucleosomes are joined by linker DNA which varies in length, resulting in differences in nucleosome repeat length (NRL). Histone H1 directly interacts with a nucleosome and is shown to influences NRL in mammals and fungi (Barra et al., 2000; Fan et al., 2005). However, how H1 influences nucleosome spacing in plants is unknown. To resolve this issue, we determined the nucleosome landscape at higher resolution by sequencing nucleosome-bound DNA fragments in wild-type and *h1* plants. We isolated mono-nucleosome-associated DNA from chromatin partially digested by Micrococcal nuclease, which digests the linker DNA.

To verify if our MNase sequencing (MNase-seq) results are comparable with published data, we initially generated average nucleosome occupancy patterns in transposons and genes in wild-type. Nucleosome positioning is conserved throughout eukaryotic organisms and Archaea (Ammar et al., 2012; Schwartz et al., 2009). Known patterns of nucleosome deposition involve the depletion of nucleosomes at the transcription start site (nucleosome-depleted region) and a highly phased nucleosome upstream of the 5' nucleosome-depleted region (Jiang and Pugh, 2009); and our three MNase-seq replicates showed this pattern in genes (Figure 3.5).

To determine the average spacing pattern of nucleosomes, we performed an autocorrelation analysis of MNase-seq data. The autocorrelation is commonly used signal processing to identify hidden periodic patterns obscured by noise. Autocorrelation was also used previously to identify the average periodicity of nucleosomes (Wan et al., 2009). We found that the nucleosome periodicity was

shorter in genes compared to transposons, indicating that NRL is longer in transposons than genes (Figure 3.6A). In addition, nucleosomes are further spaced in genes with lower expression than those with high expression levels, suggesting that gene expression intensity is negatively correlated with longer NRL (Figure 3.6B).

To further dissect H1's effects on NRL, we performed the same autocorrelation analysis in genes and transposons of *h1* mutant. We determined that the nucleosome spacing is considerably closer in *h1* compared to wild-type, with the effects being most prominent in transposons (Figures 3.6C and 3.7). Additionally, the periodicity is much shorter in genes with low expression yet this effect is essentially absent in highly transcribed genes (Figure 3.6D), likely because highly expressed genes are depleted of H1. Our results indicate that H1 increases the distance between nucleosomes in *Arabidopsis* and H1-mediated nucleosome separation is conserved across eukaryotes.

3.3.4. H1 deficiency increases DNA methylation in heterochromatin and destabilizes methylation in euchromatin

Despite its conserved structure and function in organizing chromatin structure, reduction of H1 causes contradictory changes in DNA methylation in eukaryotic organisms (Barra et al., 2000; Fan et al., 2005; Rea et al., 2012; Wierzbicki and Jerzmanowski, 2005). In order to unravel the confusion, we performed whole-genome mapping of DNA methylation in plants with mutated *H1* genes, vastly extending the dataset available from previous analyses of a smaller number of specific genes and repeats.

Our genome-wide approach revealed that *h1* has a complex DNA methylation phenotype. The average DNA methylation patterns of transposons in all sequence contexts show global increases in DNA methylation in *h1*, particularly at CG and CHG sites (Figure 3.8D). However, at the points of alignment, CHH methylation showed hypomethylation and CG and CHG methylation had relatively less hypermethylation (Figure 3.8D). This pattern suggests that DNA methylation of euchromatic and heterochromatic transposons differ in *h1*. To test this, we separated transposons based on H3K9me2 levels. We found that the heterochromatic transposons show a global increase in all cytosine contexts, while euchromatic transposons are globally hypomethylated at CHG and CHH sites and locally hypomethylated at CG sites (Figure 3.8C).

Unlike transposons, overall DNA methylation in genes was not changed (Figure 3.8B). However, the kernel density plot (Figure 3.8A) suggests local hypomethylation at specific regions of the genome in *h1*. We identified 180 loci with

significant and reproducible absolute CG methylation changes of at least 30% in *h1* plants. 61 of these DMRs are in genes, where 60/61 DMRs exhibit reduced methylation with respect to wild-type. Intergenic regions accounted for 45 DMRs with 43/45 DMRs showing hypomethylation in *h1*. The remaining 72 DMRs are in transposons with 70% (52/72 DMRs) exhibiting hypomethylation in *h1*. Also, these hypomethylated transposons exhibit low H3K9me2 (Figures 3.8E, F, G and 3.9), indicating that they are euchromatic transposons. Despite global hypermethylation in transposons, the differentially methylated regions (DMRs) are mostly hypomethylated. This is because CG methylation of heterochromatic transposons is already very high (over 70% methylation) and our threshold of 30% methylation change was too stringent to identify DMRs with hypermethylation.

3.3.5 Reduction of H1 further de-represses transposons

Histone H1 is associated with transcriptional suppression (Fan et al., 2005; Shen and Gorovsky, 1996) and a recent study showed that a reduction of H1 activates transposons in *Drosophila* (Lu et al., 2013). Since *Drosophila* lacks DNA methylation, it is possible that transposons are inactivated in *Arabidopsis h1* mutant due to DNA hypermethylation. However, mutations in some *Arabidopsis* genes de-repress transposons without affecting DNA methylation (Elmayan et al., 2005; Moissiard et al., 2012; Takeda et al., 2004; Vaillant et al., 2006), indicating that the transposons in *h1* could be activated in spite of hypermethylation. To resolve this question, we generated a map of RNA in *h1* and wild-type. The RNA-seq data revealed that the vast majority of transposons remain inactive in *h1*. We detected only three transposons that are silent in wild-type and significantly expressed in h1, and identified seven repressed transposons in *h1* compared to wild-type (Table 3.1). These over-expressed and repressed transposons are hypomethylated and hypermethylated in *h1*, respectively (Figure 3.10 and Table 3.1). Therefore, our data indicate that H1 affects transposon expression in *Arabidopsis* largely via DNA methylation.

3.4 Discussion

3.4.1 The role of H1 in genes

Previous studies indicate that H1 binding influences transcription factor binding (Juan et al., 1994; 1997), as well as transcription initiation and elongation (Cheung

et al., 2002; O'Neill et al., 1995). Our findings regarding the distribution of H1 reflect previous studies on the role of H1 in transcription machinery regulation. We identified that H1 is present in *Arabidopsis* genes and that H1 abundance is anticorrelated with transcription (Figure 3.2A and B). More specifically, the less expressed genes had more H1 at their promoters, transcription start sites, gene bodies, and 3' transcription end sites compared to highly expressed genes (Figure 3.2A and B). This binding pattern suggests that H1 hinders binding of transcription machinery at the promoter and prevents Pol II transcription initiation and elongation.

A link between H1 and DNA methylation was proposed based on their co-localization (Ball et al., 1983). However, it is controversial whether H1 has a higher affinity to methylated DNA than it does to unmethylated DNA. Campoy et al. identified that H1 has the same affinities for methylated and non-methylated naked DNA (Campoy et al., 1995). At the same time, McArthur et al. determined that H1 preferentially binds to methylated oligonucleotides compared to unmethylated DNA (McArthur and Thomas, 1996). Our *in vivo* analyses show that H1 is preferentially associated with methylated genes compared to unmethylated genes (Figure 3.3). This enrichment of H1 is not due to the static silent nature of a methylated element such as transposons, because methylated genes are shown to be highly expressed (Zilberman et al., 2007). To further investigate this issue, we are currently mapping H1 proteins in the *met1* mutant, where global reduction of CG methylation is observed, including in genes (Cokus et al., 2008; Lister et al., 2008). We anticipate observation of a reduction of H1 enrichment in genes that lose DNA methylation in *met1*.

3.4.2 H1 inhibits DNA accessibility

Our data indicates that the role of H1 is not to specifically inhibit or promote DNA methylation, but rather to inhibit all nucleoproteins from binding DNA (Figure 3.8). H1 may render the nucleosome-bound DNA less accessible to euchromatin factors, such as DNA glycosylases (Choi et al., 2002; Gong et al., 2002; Penterman et al., 2007), histone acetyltransferases (Probst et al., 2004; Qian et al., 2012) and other enzymes that counteract DNA methylation (Deleris et al., 2010; Searle et al., 2010). In heterochromatin, H1 may prevent accessibility of heterochromatic proteins such as DNA methyltransferases (Law and Jacobsen, 2010), and inhibits hypermethylation. Our results suggest that the balance between exclusion and access mediated by H1 is essential for the stable propagation of chromatin states.

Moreover, the complex DNA methylation defects in *h1* provide a plausible explanation for the observations that H1 reduction can cause both localized losses of

DNA methylation in mammals and gains of methylation in *Ascobolus* (Barra et al., 2000; Fan et al., 2005). Therefore, performing genome-wide analyses in mutants of these organisms that lack H1 is expected to provide clues about the conserved relationship between H1 and DNA methylation in eukaryotes.

3.5 Materials and methods

Chromatin affinity purification sequencing (ChAP-seq)

ChAP experiements were performed according to (Zilberman et al., 2008) with minor modifications. The N-terminus of *H1.1* and *H1.2* was tagged with biotin ligase recognition peptide (BLRP) and the constructs were driven by the endogenous promoters. The transgenic plants were grown in liquid culture for 4 weeks. The roots were vacuum infiltrated in 1% formadehyde solution for 15 minutes to crosslink the chromatin then vacuum infiltrated for 5 minutes in 200mM glycine to stop the reaction. 1g of roots were ground in liquid nitrogen to fine powder and 15mL lysis buffer containing 50mM Tris pH 7.5, 150 mM NaCl, 5mM $MgCl_2$, 0.1% SDS, 0.1% NP-40, 1mM PMSF was added. The suspension was incubated on ice for 30 minutes then separated into 1 mL aliquots. The root suspension was fragmented by sonication to 0.5-2 kb, and then centrifuged to pellet debris. 90% of sheared chromatin containing supernatant was incubated with 200 ul of High Capacity Streptavidin Agarose beads (Thermo Scientific) for 1 hour at 4ºC, and 10% of suspension was saved for use as an input control (IC). The agarose beads were washed with TNE buffer containing 50mM Tris pH 7.4, 100mM NaCl, 0.1mM EDTA five times then resuspended in 400ul TNE. The affinity purified (AP) sample and the IC were subjected to reverse cross-linking with 200 mM NaCl at 65ºC overnight. The samples were incubated with 1 ul RNAse A for 30 minutes and then 4 ul Proteinase K with 0.5% SDS for 2 hrs at 50ºC. DNA from the reactions were purified by phenol-chloroform extraction. Genomic DNA libraries were constructed with the AP- and IC-DNA according to the manufacturer's instructions (Illumina).

Biological materials

h1.1 (*AT1G06760*, SALK_128430C) and *h1.2* (*AT2G30620*, GABI_406H11) T-DNA lines were obtained from ABRC and GABI-Kat collections, respectively. T-DNA insertions were confirmed by PCR-based genotyping. The same *h1* line was previously used in (Rea et al., 2012). Roots were grown under continuous light in a sterile liquid medium for 4 weeks, and seedlings were grown under continuous light for 10 days after two days of cold treatment. All mutants are in the Columbia ecotype.

Quantitative PCR

We used three biological replicates of *h1* and wild-type roots. RNA was extracted using the RNeasy Plant Mini Kit (Qiagen). Genomic DNA was digested on column using the RNA-Free DNase Set (Qiagen). cDNA libraries were generated using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific), and the primers spanning the exons were used to determine the transcription level.

Western blot and Coomassie staining

Antibodies were produced by Pro-Sci against synthetic peptides, then affinity purified using the appropriate epitopes as described previously (Ascenzi and Gantt, 1997). 1-2g of roots were ground into a fine powder and lysed in a solution containing 10 mM EDTA, 0.12 M Tris-HCl, 4% SDS, 10% beta-mercaptoethanol, 5% glycerol, and 0.005% bromophenol blue. Proteins were separated in an Any-KD mini PROTEAN TGX pecast gel (Bio-rad). The proteins ran at 100 V for 1.5 hrs at room temperature.

For Coomassie staining, the gel was stained in a colloidal Coomassie solution containing 0.02% (w/v) of Coomassie Brilliant Blue G-250, 5% (w/v) aluminum sulfate, 10% (v/v) ethanol (96%), 2% (v/v) orthophsphoric acid (85%) as previously described (ref). The gel was incubated with Milli-Q water for 10 minutes three times. Then, the gel was stained in the Coomassie solution overnight with a gentle shake. After incubation, the gel was washed twice with Milli-Q water.

For Western blotting, proteins were transferred onto nitrocellulose in a mini-PROTEAN tetra blotting module according to the manufacturer's instructions (Bio-Rad). Binding of antibodies was performed using a 1:5,000 dilution of primary antibody and a 1:10,000 dilution of HRP-conjugated goat antirabbit immunoglobin G secondary antibody (Thermo Scientific). Signal was detected using SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

Microccocal nuclease sequencing (MNase-seq)

1 g of roots was ground into a fine powder in liquid nitrogen, resuspended in 20mL of HBM buffer containing 20mM Tris-HCl pH 7.6, 0.44M sucrose, 10mM $MgCl_2$, 2mM spermine, and 0.1% Triton X-100, then filtered through two layers of miracloth into a 30mL round bottom glass tube. The homogenate was centrifuged at 2,000g at 4ºC for 10 minutes and the pellet was washed in 1 mL HBB buffer containing 25 mM Tris-HCl pH 7.6, 0.44M sucrose, 10mM $MgCl_2$ and 0.1% Triton X-100. Nuclei were further spun down at 200 g at 4ºC for 2 minutes. Then, the pellet

was resuspended in 1 mL of TNE buffer containing 10mM Tris-HCl pH 8.0, 100mM NaCl, and 1mM EDTA. 4 ul of 1 M $CaCl_2$ was added prior to digestion. Typically 1 g of *Arabidopsis* roots yielded approximately 400 ul of nuclei after settling by gravity for 15 minutes. 20 ul of nuclei were saved as a non-MNase-treated control. The rest of the nuclei were aliquoted into 100 ul volumes and then brought to 1 mL using TNE. 1 ul of the diluted MNase Stock (1/20 dilution of 200 u/ml stock, Sigma) was added per aliquot and the samples were incubated at 37ºC for 1, 3, 5, and 7 minutes under a gentle shake. The reactions were stopped by the addition of 50 ul of 0.5 M EDTA. Digested nuclei were centrifuged at maximum speed for 5 minutes at 4ºC, and released soluble nucleosomes were collected from the supernatant. The supernatant was RNase A treated for 10 minutes at room temperature and then Proteinase K treated with 0.5% SDS at 55ºC for 10 minutes. Purified DNA samples including the control were run on a 2% agarose gel. The mononucleosome corresponding bands (~150 bp) from the samples with most enriched intact mononucleosomes were gel purified (Qiagen). Illumina libraries were constructed and sequenced to generate paired-end 36 base reads.

Bisulfite sequencing (BS-seq)

About 500ng of genomic DNA was isolated from roots, fragmented by sonication, end-repaired, and ligated to custom synthesized methylated adapters (Eurofins MWG Operon) according to the manufacturer's instructions (Illumina) for gDNA library construction. Adapter-ligated libraries were subjected to two successive treatments of sodium bisulfite conversion using the EpiTect Bisulfite kit (Qiagen) as outlined in the manufacturer's instructions. The bisulfite-converted libraries were then amplified by PCR using the following conditions: 2.5 U of ExTaq DNA polymerase (Takara Bio), 5 ul of 10x ExTaq reaction buffer, 25 uM dNTPs, 1 ul universal primer, and 1 ul indexed primer (50ul final). PCR reactions were carried out as follows: 95ºC for 3 minutes, then 12-14 cycles of 95ºC for 30 seconds, 65ºC for 30 seconds, and 72ºC for 60 seconds. The enriched libraries were gel purified (~300 bp band) prior to quantification with a Bioanalyser (Agilent). Sequencing on the Illumina platform was at the Vincent J. Coates Genomic Sequencing Laboratory at UC Berkeley.

mRNA purification and sequencing (RNA-seq)

Total RNA samples from 4-week roots were isolated using the RNeasy Mini Kit (Qiagen) including on-column DNase treatment. mRNA was purified from 10-50 ug of total RNA by two cycles of poly-A enrichment using the Oligotex kit (Qiagen) followed by an rRNA removal step using the RiboMinus Plant Kit for RNA-seq (Invitrogen). Purified mRNA samples were eluted with 9 ul RNase free water and fragmented with 1 ug of 10x fragmentation buffer (Ambion) at 70ºC. Reactions were

stopped after 5 minutes by adding 1 ul Stop Buffer, and RNA was purified by ethanol precipitation. cDNA was synthesized from 100-300 ng of mRNA using SuperScript III reverse transcriptase (Invitrogen). Double-stranded DNA was synthesized by SuperScript Double Stranded cDNA Synthesis Kit according to the manufacturer's instruction (Invitrogen). DNA was cleaned with a QIAquick PCR spin column (Qiagen), and sequencing adapters were ligated according to the Illumina protocol. The Illumina library was amplified for 18 cycles using Phusion DNA polymerase (NEB). Bands around 300 bp were gel purified and cloned for validation. Traditional sequencing confirmed that the libraries were properly constructed showing high percentage of mRNA over rRNA. The libraries were sequenced on GAII Illumina sequencer to generate 37 base reads.

3. 6. Genomic data analyses

Gene and TE average pattern analysis of sequencing data

Arabidopsis genes or transposons were aligned at the 5' or the 3' end. To avoid averaging the edges of short genes/transposons and the middles of long sequences, we removed transposons smaller than 250 bp, and genes smaller than 2000 bp in our H1 enrichment analysis and genes smaller than 1500 bp for other analyses. Genes with CG methylation over 60% were excluded from the analysis because they behave like transposons (Zemach et al., 2013). Genes with low CG methylation (<20%) were also excluded because they decrease the dynamic range without substantively contributing to the analysis.

Autocorrelation analysis

Single base autocorrelation plots were generated using nucleosome center positions using the STATA statistical program (reference or version). The first three peak positions, estimated with Gaussian functions using Fityk 0.9.8, were identified to determine the average nucleosome repeat length (NRL).

Kernel density plots

The density plots were generated with the difference between *h1* and wild-type roots in 50 bp windows. We plotted windows with at least 10 informative sequenced cytosines and fractional methylation of at least 30% for CG and 10% for CHG and CHH in *h1* or wild-type. Genes with over 60% and under 20% CG methylation were excluded from the analysis. H3K9me2 and H3 data were obtained from (Moissiard

et al., 2012). Transposons overlapping windows with $\log_2$(H3K9me2/H3) smaller than -1 and larger than 1 were considered euchromatic and heterochromatic, respectively.

Identification of differentially methylated regions (DMRs)

Methylated CG difference was initially compared between *h1* and wild-type roots. Adjacent 50 bp windows with a fractional CG difference of at least 0.3 were merged, allowing up to 100 bp gaps without coverage. We identified DMRs of at least 250 bp with CG methylation difference > 0.3 (positive DMRs) or < -0.3 (negative DMRs) and Fisher's exact test p-value $<10^{-6}$ in roots and seedlings.

Identification of transposons with altered expression

We identified annotated transposons exhibiting CG and CHG methylation in wild-type or *h1* (CG methylation only is a hallmark of genes) with an expression change of at least 2 fold in *h1* compared to wild-type and Fisher's exact test p-value $<10^{-4}$.

3.6 Figures and tables

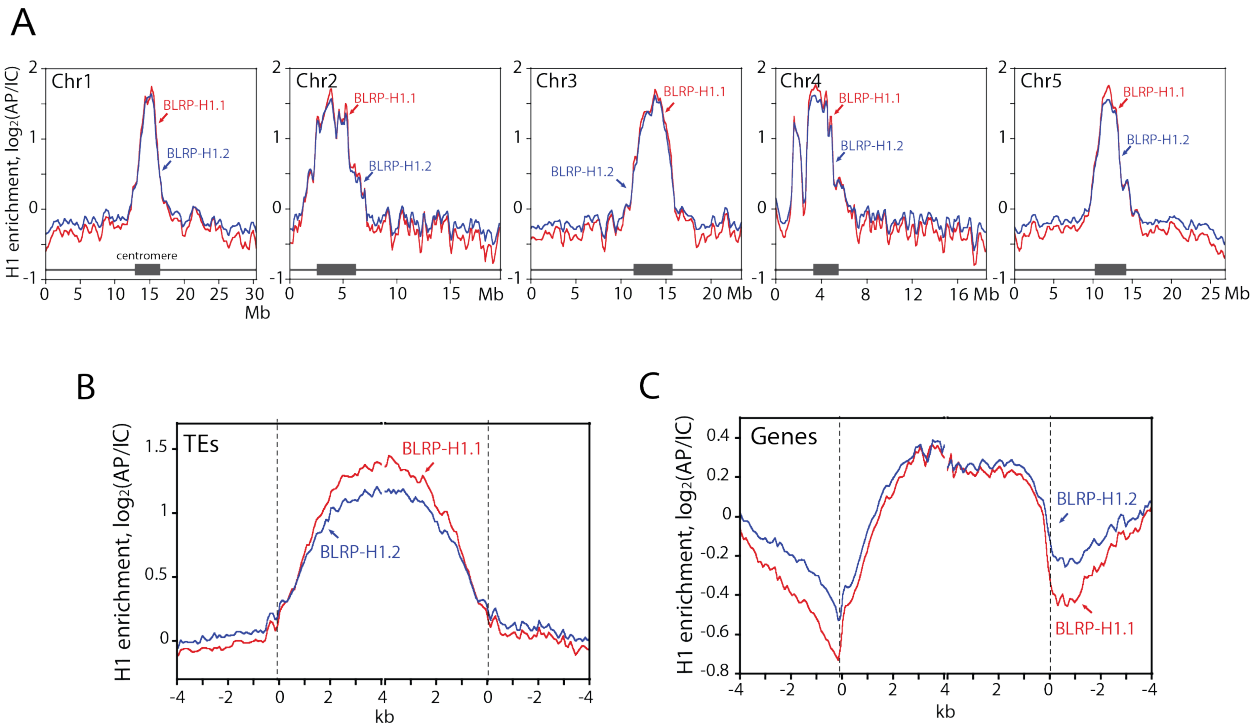

Figure 3.1. H1 is enriched in transposons. (A) H1 enrichment scores in 200 kb windows are plotted for each chromosome. The gray boxes represent pericentric heterochromatin. Transposons (B) and genes (C) were aligned at the 5' and the 3' end (dashed lines, points of alignment), and average H1 enrichment for each 5 bp interval is plotted. The red traces represent H1.1 and the blue traces represent H1.2 from ChAP-seq.

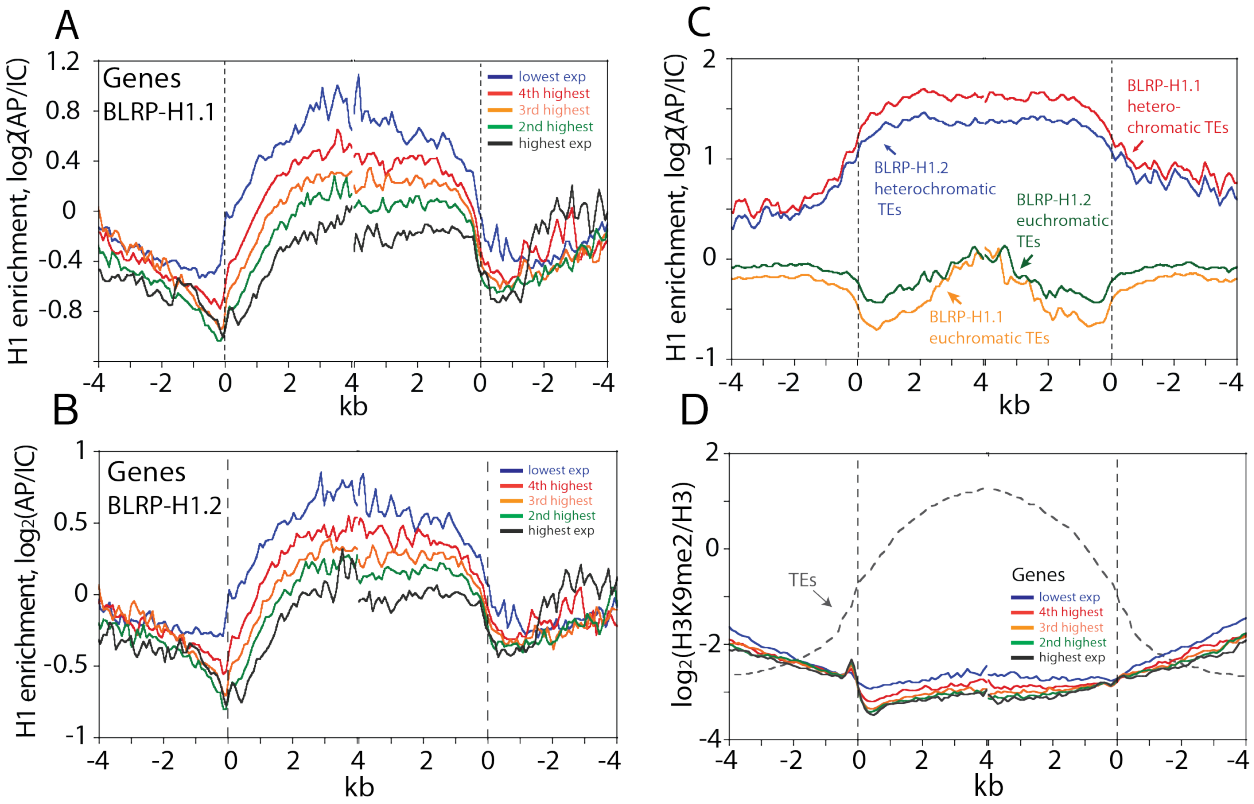

Figure 3.2. H1 is enriched in inactive chromatin. (A-D) Genes (A, B, D) or transposons (C) were aligned as Figure 3. 1B and average H1 enrichment (A-C) or H3K9me2 (D) for each 5 bp interval is plotted. The dashed lines represent the points of alignment. Genes were grouped into quintiles by expression level in (A, B, D). Heterochromatic and euchromatic transposons shown in (C) were categorized based on H3K9me2 enrichment (ref).

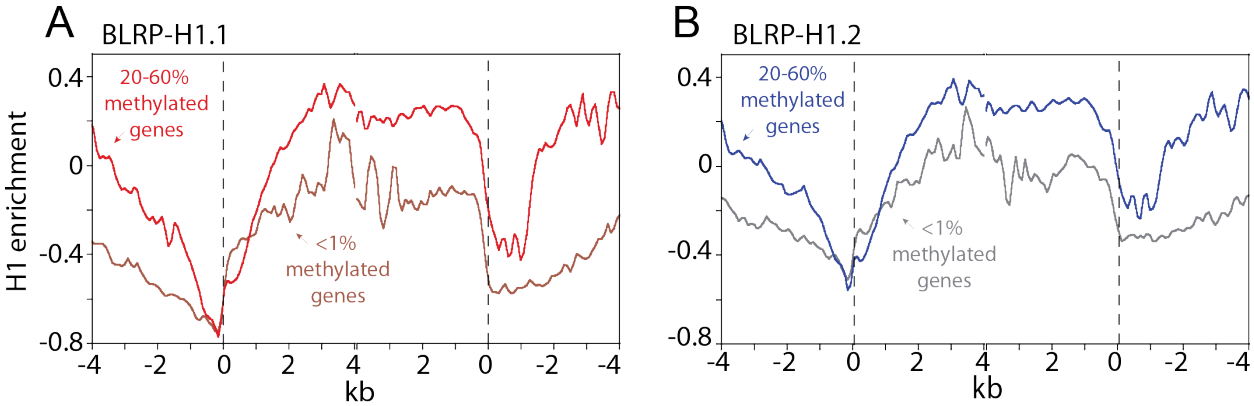

Figure 3.3. Methylated genes are enriched with H1. H1.1 (A) and H1.2 (B) enrichment is plotted for methylated (20-60% average methylation) or unmethylated (less than 1% average methylation) genes as in Figure 3. 1B.

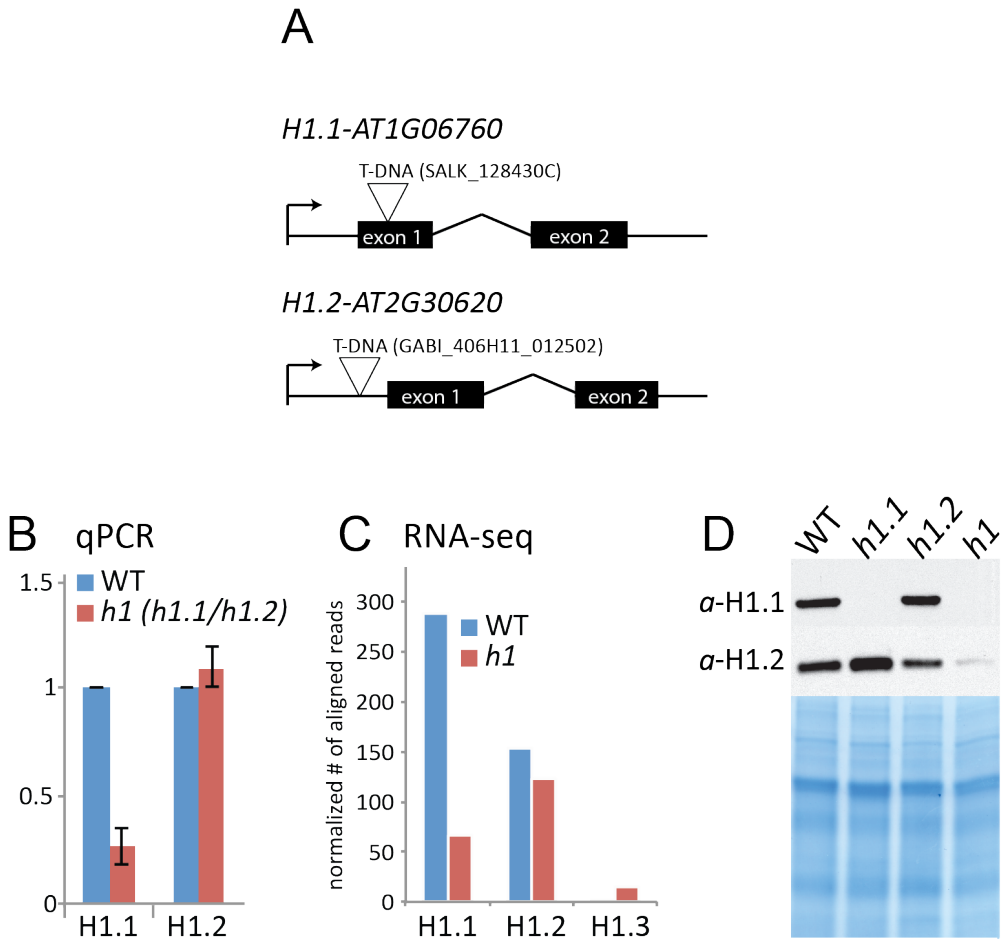

Figure 3.4. *h1* mutant characterization. (A) Genomic structure of *H1.1* and *H1.2*, and their T-DNA insertion locations (empty triangle) are shown. Black lines represent UTRs, kinked lines represent introns, and black boxes represent exons. (B-C) qPCR analysis (B) and normalized reads from RNA-seq. (D) Western blot of wild-type, *h1.1* (only in Western), *h1.2* (only in Western), and *h1*.

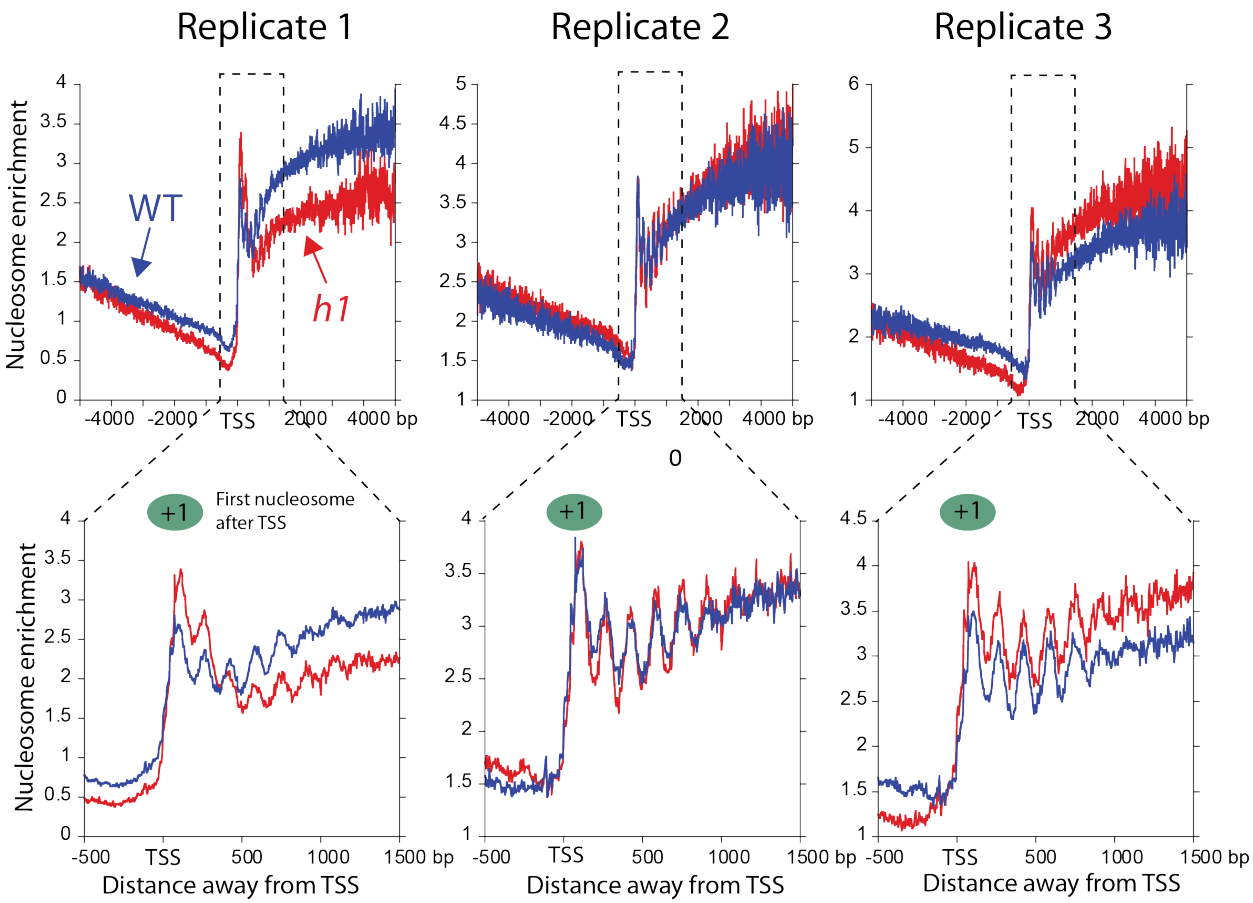

Figure 3.5. Nucleosome deposition in genes. Wild-type (WT, blue trace) and *h1* mutant (red trace) of nucleosome enrichment in genes. Genes were aligned as in Figure 3. 1C and average nucleosome levels of each 5 bp window interval are plotted. The green oval represents the highly phased nucleosome just upstream of transcription start site (TSS).

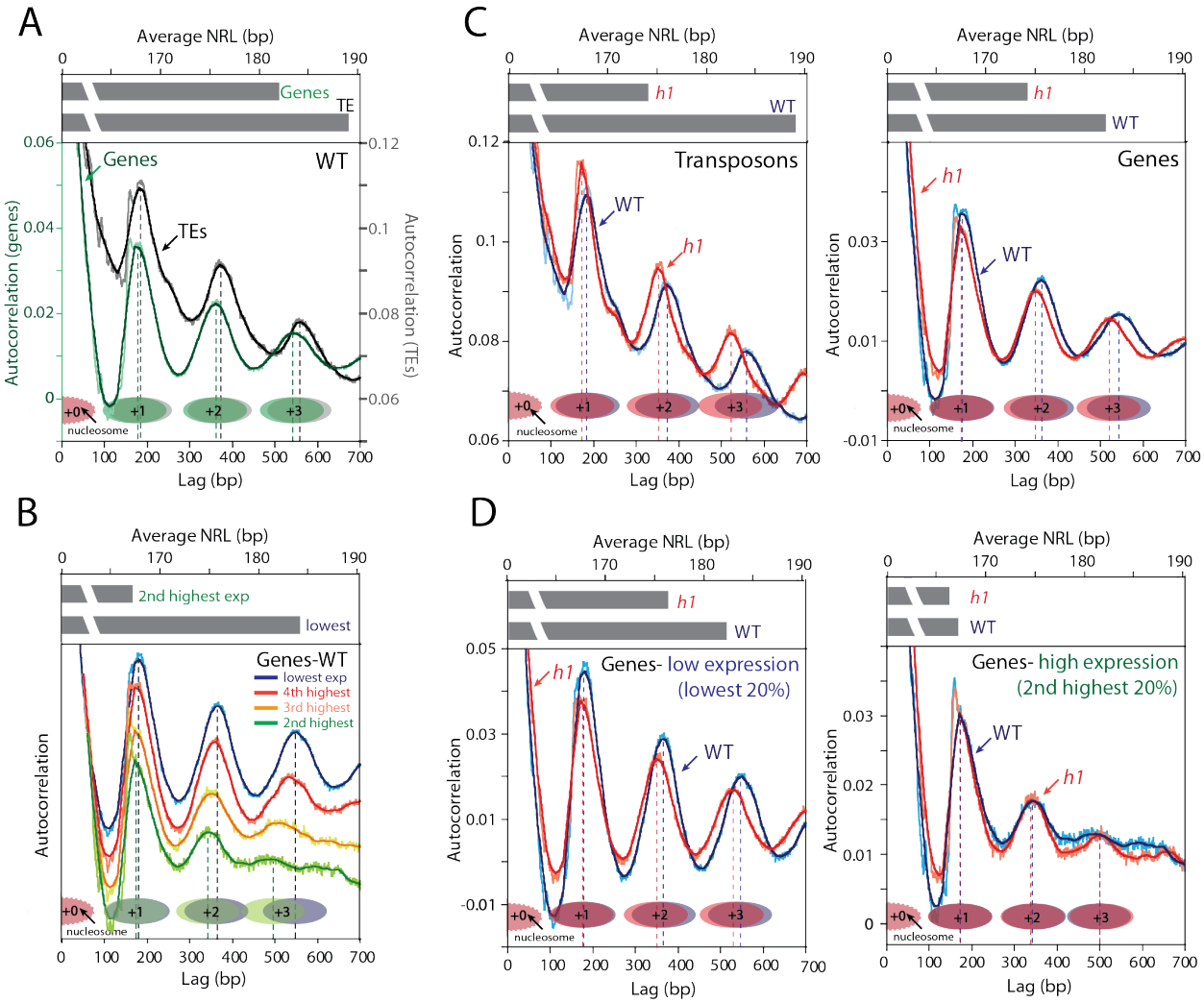

Figure 3.6. H1 increases the distance between nucleosomes. (A-D) Single base autocorrelation plots, which show the predominant periodicity within the data, of nucleosome center positions from three averaged biological replicates of wild-type or *h1* plants. Gray bars above the plots show the average nucleosome repeat length (NRL). Individual replicates are shown in Figure 3. 6. The transcriptional quintiles shown in (B) and (D) are the same as Figure 3. 2A; the highest expression group is not shown due to weak periodicity. Note that h1 causes substantial NRL shortening in transposons and weakly expressed genes, but not in genes with high expression.

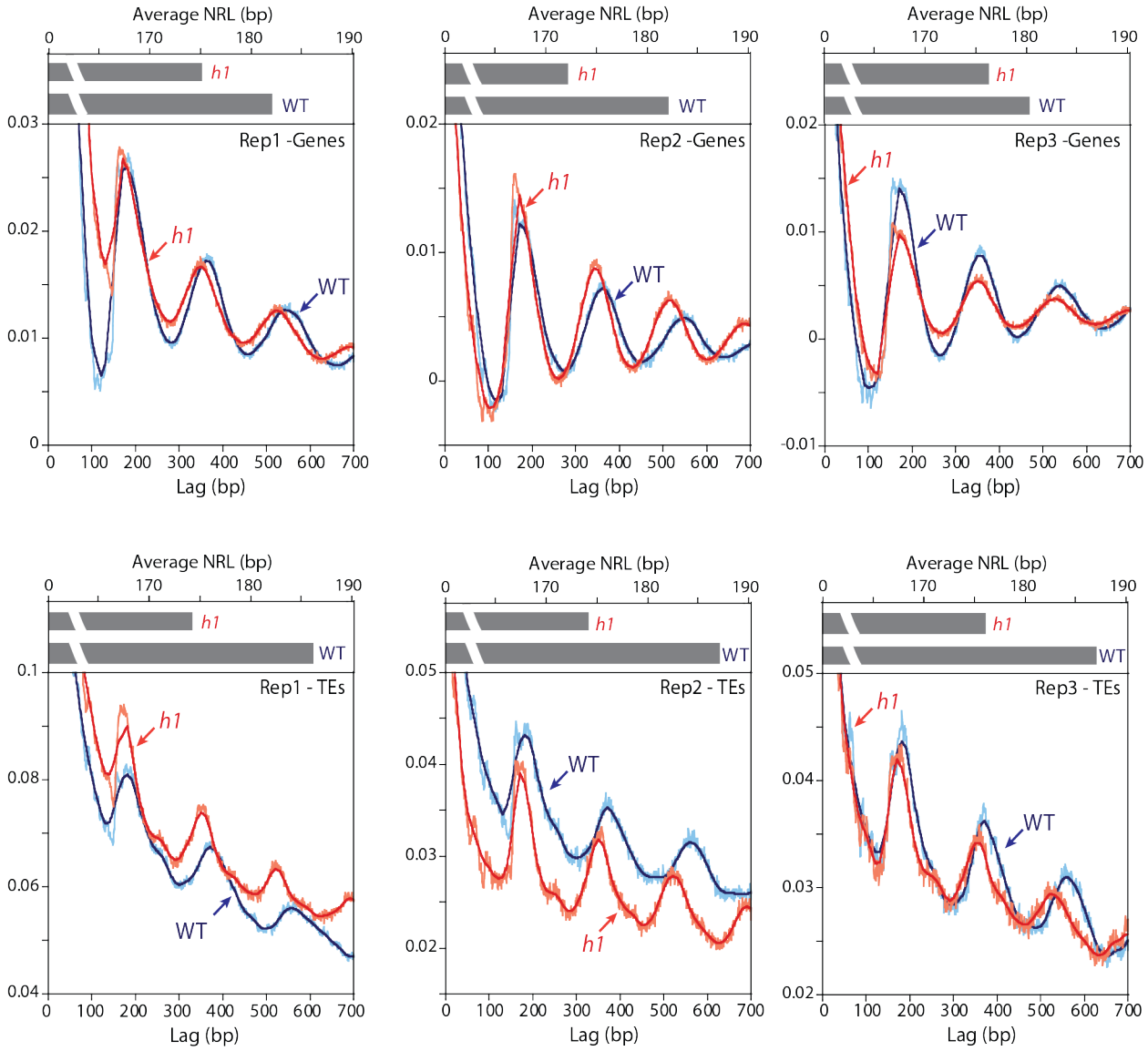

Figure 3.7. Autocorrelation analysis of three biological replicates. Autocorrelation plots show the nucleosome periodicity within genes and transposons in three wild-type (blue trace) and *h1* (red trace) biological replicates. All replicates show that H1 increases the distance between nucleosomes.

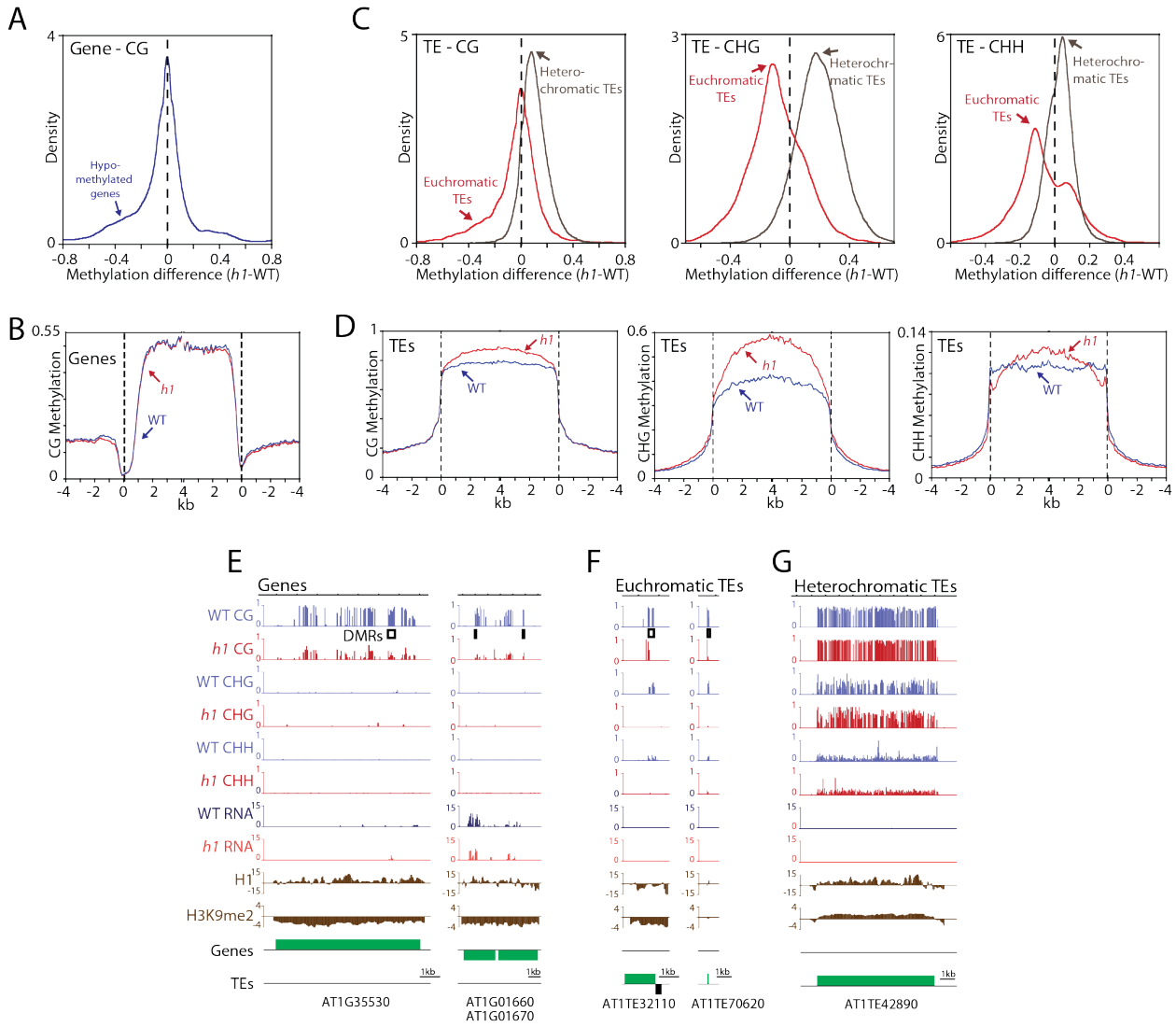

Figure 3.8. H1 loss increases DNA methylation of heterochromatin and destabilizes methylation in euchromatin. (A, C) Kernel density plots of methylation differences between *h1* and wild-type (positive numbers indicate greater methylation in *h1*). Heterochromatic and euchromatic transposons are categorized as in Figure 3. 2C. The arrows emphasize global differences (a shifted peak) or extensive local difference (a broad shoulder). (B, D) Genes and transposons were aligned as in Figure 3. 1B and C and average methylation levels of each 100 bp window interval are plotted. (E-G) DNA methylation (CG, CHG, CHH), RNA, H1, and H3K9me2 levels of gene (E), euchromatic transposon (F), and a heterochromatic transposon (G) are emphasized in green. Black boxes represent CG DMRs.

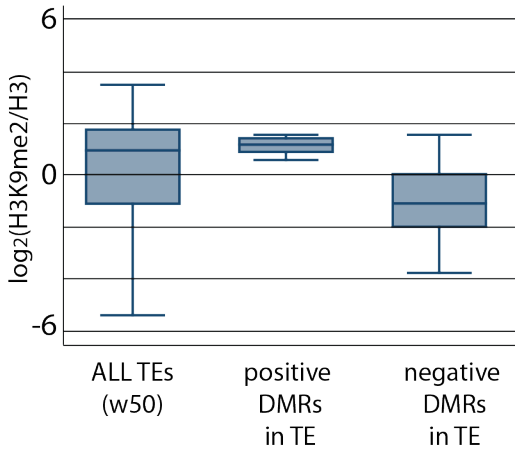

Figure 3.9. Hypomethylated transposons in *h1* are euchromatic. Box plots of H3K9me2 levels in 50 bp windows within hypermethylated (positive DMR) transposon DMRs and hypomethylated (negative DMR) transposon DMRs in *h1*. All 50 bp windows within transposons are plotted as a control. Each box encloses the middle 50% of the distribution, with the horizontal line marking the median, and vertical lines marking the minimum and maximum values that fall within 1.5 times the height of the box.

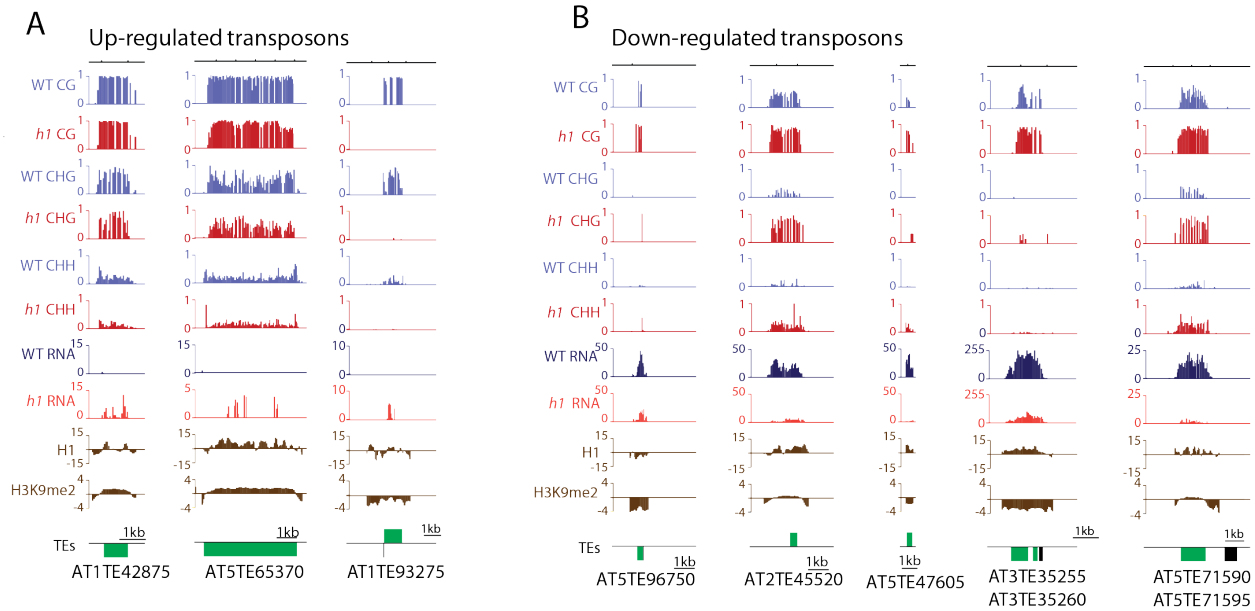

Figure 3.10. Alteration of transposon expression in *h1* depends on DNA methylation. (A-B) Examples of upregulated transposons (A) and down-regulated transposons (B).

TEs with altered expression

Chr	Start	End	ID	type	WT_RNA	*h1*_RNA	RNA diff (*h1*-WT)	p-value	mCG diff (*h1*-WT)	mCHG diff (*h1*-WT)	mCHH diff (*h1*-WT)
Down regulated TEs											
chr2	10630652	10631003	AT2TE45520	ATHILA6A	141.624	30.2323	-111.3917	4.14E-09	0.35	0.70	0.11
chr3	8410620	8411277	AT3TE35255	ATHILA6A	2254.97	529.066	-1725.904	1.45E-134	0.19	0.07	0.01
chr3	8411467	8411641	AT3TE35260	VANDAL1	752.982	188.822	-564.16	1.99E-41	0.56	0.00	0.01
chr5	13451924	13452252	AT5TE47605	ATHILA6A	164.744	3.90935	-160.8346	5.40E-26	0.39	0.29	0.11
chr5	19908422	19909072	AT5TE71590	ATHILA6A	197.049	21.8924	-175.1566	6.42E-22	0.33	0.51	0.28
chr5	19909089	19909745	AT5TE71595	ATHILA6A	201.642	14.8555	-186.7865	8.23E-24	0.35	0.53	0.18
chr5	26875287	26875627	AT5TE96750	ATENSPM6	341.583	83.5298	-258.0532	1.45E-16	0.11	0.01	0.00
Upregulated TEs											
chr1	13127091	13127992	AT1TE42875	ATSINE2A	0	47.8244	47.8244	8.24E-10	0.01	-0.02	-0.08
chr1	28518723	28519779	AT1TE93275	HELITRON1	0	22.9349	22.9349	6.95E-05	-0.93	-0.43	-0.11
chr5	18159383	18164125	AT5TE65370	ATCOPIA21	0	26.1927	26.1927	1.64E-05	-0.05	-0.11	-0.07

Table 3.1. Transposons with altered expression in *h1* mutant. List of identified transposons with altered expression. Note that hypermethylation (positive methylation differences, last three columns) leads to down-regulation of expression and hypomethylation (negative methylation differences, last three columns) leads to overexpression.

Chapter 4:

The nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin

Part of the following chapter has been published as a peer reviewed article in *Cell*. My contribution to this article was determining the relationship between DDM1 and H1.

Zemach A*, Kim MY*, Hsieh PH*, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, and Zilberman D (2013) "The nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin." Cell, 153:193-205.
*These authors contributed equally.

4.1 Abstract

DDM1 is a chromatin remodeler necessary for DNA methylation in heterochromatin. Loss of DDM1 significantly reduces DNA methylation levels, suggesting that DNA methyltransferases cannot access their targets without DDM1. In contrast, H1 inhibits the DNA methyltransferase accessibility of DDM1 target loci. The link between DDM1 and H1 is apparent, but the functional relationship between these two chromatin components is unknown. Here, we show that DDM1 is required to overcome the DNA inaccessibility mediated by H1 for DNA methylation. H1 loss largely suppresses the profound genomic hypomethylation caused by mutation of the nucleosome remodeler DDM1 and eliminates the preferential dependence of heterochromatic transposons on DDM1. Moreover, loss of H1 suppresses the rate of onset of developmental abnormalities. Our results demonstrate that H1 stabilizes epigenetic inheritance and creates a less accessible chromatin configuration that requires DDM1 for DNA methylation.

4.2 Introduction

DDM1, the Snf2 family nucleosome remodeler, is essential for normal DNA methylation (Jeddeloh et al., 1999; Zemach et al., 2013). DDM1 can shift nucleosomes *in vitro* (Brzeski and Jerzmanowski, 2003) and its mutation has been reported to cause a profound loss of DNA methylation from some transposons and repeats (Zemach et al., 2013). Mutation of Lsh, the mouse homolog of DDM1, causes a similar methylation phenotype (Tao et al., 2011), indicating that DDM1 remodelers are an ancient component of the DNA methylation pathway. The loss of DDM1 leads to strong transcriptional activation of transposons (Lippman et al., 2004), and inbred *ddm1* mutant lines have pleiotropic phenotypes (Kakutani et al., 1996).

A previous study proposed that DDM1 functions to provide methyltransferases access to the DNA (Zemach et al., 2013), and our *h1* data (see Chapter 3) suggest that H1 is one of the restrictive factors overcome by DDM1. Consistent with this idea, loss of methylation in *ddm1* mutants is correlated with wild-type H1 levels (Figure 4.3). To directly test our hypothesis, we crossed heterozygous *ddm1*/+ plants with homozygous *h1* plants to generate *h1ddm1*, as well as control *h1*, *ddm1* and wild-type siblings. Here, we report genome-wide analysis of DNA methylation of the *h1ddm1* double mutant. Our data suggest that DDM1 is specialized for remodeling heterochromatic, H1-bound nucleosomes to allow DNA methyltransferases, and likely other proteins, access to the DNA.

4.3 Results

4.3.1 Histone H1 mediates the dependence of heterochromatic DNA methylation on DDM1

Initially, we investigated whether the DNA methylation of controls (*h1, ddm1*, and wild-type) from the *h1ddm1* cross are comparable with the previously described results from the uncrossed line. Homozygous *h1* seedlings derived from the cross exhibit methylation defects very similar in DNA methylation in all cytosine contexts to that of *h1* roots from our study (Figures 3.8 and 4.1). The general patterns of CHH methylation in transposons are similar in both samples with short euchromatic transposons being less methylated and heterochromatic transposons being more methylated. However, the overall DNA methylation level is lower in *h1* seedlings than in *h1* roots in all cytosine contexts (Figures 3.8D and 4.1). This difference is likely explained by age-dependence, as older tissues show more total methylation than younger tissues in animals and plants (Bellizzi et al., 2012; Li et al., 2010; Maegawa et al., 2010; Zemach et al., 2010).

Loss of H1 almost completely suppresses the reduction of transposon CHH methylation in *ddm1* and greatly ameliorates the reduction of transposon CG and CHG methylation (Figures 4.1A and 4.2). Most strikingly, H3K9me2-rich transposons are not preferentially hypomethylated in *h1ddm1* plants, as they are in *ddm1* (Zemach et al., 2013) (Figures 4.2). Instead, *h1ddm1* causes H3K9me2-rich transposons to lose less DNA methylation than more euchromatic transposons (Figure 4.2A, blue trace), similar to *h1* (Figure 4.2, red trace), indicating that the loss of DDM1 affects transposons similarly regardless of H3K9me2 levels when H1 is not present (Figure 4.2B). Lack of H1 still destabilizes the methylation of euchromatic transposons when combined with *ddm1*, but heterochromatic transposons are methylated rather efficiently when both DDM1 and H1 are absent (Figures 4.2A and 4.4). The rescue of the *ddm1* phenotype requires a substantial loss of H1, as the methylation patterns in *h1.1ddm1* and *h1.2ddm1* closely resemble that of *ddm1* (Figure 4.1B). Also, CHH methylation in *h1ddm1* is similar to that in *h1* at *Gypsy* and MuDR elements and, to a lesser extent, at *Copia* and LINE elements (Figure 4.5). Our results indicate that the differential importance of DDM1 for the maintenance of DNA methylation in heterochromatic versus euchromatic transposons is governed by H1 (Figure 4.3). The similar phenotypes of *ddm1* and *lsh* mutants (Tao et al., 2011) strongly suggest that our conclusions apply to mammals and other animals.

4.3.2 Developmental defects caused by loss of DDM1 are suppressed by loss of H1

We identified that the reduction of H1 greatly ameliorates the effects of a loss of DNA methylation in *ddm1*. This result led us to determine whether reduction of H1 also suppresses the appearance of abnormal phenotypes that result from the *ddm1* mutation (Kakutani et al., 1996). We allowed the plants from the cross to self-fertilize once and investigated morphological defects across the F3 population of *h1*, *ddm1*, *h1ddm1*, and wild-type. Our *ddm1* plants showed pleiotropic phenotypes, varying from wild-type like phenotypes to very severe mutant phenotypes (Figure 4.6A). The characteristics of mutant phenotypes included reduced apical dominance, small leaves, late flowering, and reduced fertility, as previously described (Kakutani et al., 1996). The *h1ddm1* plants resembled wild-type and *h1* plants (Figure 4.6B), indicating that reduction of H1 rescues the mutant phenotypes caused by the *ddm1* homozygous mutation.

4.4 Discussion

4.4.1 DNA methyltransferases require DDM1 to overcome DNA accessibility inhibition by H1

DDM1 is required for methylation in all sequence contexts of highly heterochromatic transposons (Figure 4.1) (Zemach et al., 2013). This requirement is reduced at less heterochromatic elements, is least in euchromatic genes (Figure 4.2), and depends on histone H1 (Figure 4.3). Lack of access to DNA is postulated to be a core property of heterochromatin that enforces gene silencing by preventing binding of transcription factors and RNA polymerase (Grewal and Jia, 2007). At the same time, stable maintenance of inaccessible heterochromatin requires DNA methylation and histone modifications like H3K9me2 that are catalyzed by enzymes that need to access chromatin. This conundrum is exemplified by the RdDM pathway, which silences transposon expression yet requires transposon transcripts from Pol IV and Pol V to function (Haag and Pikaard, 2011). Our results indicate that H1 restricts access to nucleosomal DNA and that DDM1 overcomes this restriction to enable the maintenance of DNA methylation and silencing of diverse transposons. Without DDM1, DNA methyltransferases cannot efficiently methylate inaccessible heterochromatic transposons (Figure 4.1), leading to de-repression and transposition (Zemach et al., 2013). Without H1, less heterochromatic sequences lose methylation (Figure 3.8 and 4.2A), presumably because they become more

accessible to enzymes that catalyze euchromatic histone modifications and antagonize DNA methylation (Ibarra et al., 2012; Probst et al., 2004; Qian et al., 2012).

4.4.2 Future directions—transcriptome analysis and transgenerational DNA methylation analysis in *h1ddm1*

Reduction of H1 partially restores DNA methylation loss in *ddm1* and eliminates the *ddm1*-like morphological defects. Additionally, inbred *ddm1* shows increased rate of transposition (Tsukahara et al., 2009). Therefore, it will be interesting to determine whether the partial DNA methylation restoration in *h1ddm1* also suppresses transposon activation.

Moreover, *ddm1* shows an accumulation of DNA methylation defects in genes and a continuous decrease in overall transposon DNA methylation (Kakutani et al., 1999; Tsukahara et al., 2009). Because loss of H1 did not fully restore the loss of DNA methylation in *ddm1*, it is curious how *h1ddm1* DNA methylation patterns change after several generations of selfing. Would *h1ddm1* transgenerational DNA methylation defects be similar to *ddm1*? Or would the loss of H1 fully restore DNA methylation in the later generations? These questions require further investigation.

4.5 Materials and methods

Biological materials

The *ddm1* mutant line (*AT5G66750, ddm1-2*) was described previously (Jeddeloh et al., 1999). The *ddm1* was extensively backcrossed before crossing to remove any unlinked mutation caused by a homozygous ddm1 mutation. *h1* mutant (*h1.1/h1.2*) generation and T-DNA insertion confirmation is described in Chapter 3. PCR and Sanger sequencing confirmed the point mutation in ddm1.

Bisulfite sequencing (BS-seq)

The libraries were constructed as described in Chapter 3 with minor modifications. The DNA was purified with the solid-phase reversible immobilization method using Agencourt AMPure XP beads throughout the construction. Before the library construction, DNA size selection was performed by “double-sided AMPure”

by purifying the sonicated DNA with 0.55x beads and then with 1x beads. This size selection removed very long DNA and short DNA (<150 bp), leaving 150-300 bp DNA fragments for the library construction. The final library was purified with 0.8x AMPure XP beads twice.

4.6. Genomic data analyses

Curve fits of log_2 DNA methylation ratios

M-spline fit of log_2 ratios and the absolute level of DNA methylation in 50 bp windows was plotted against log_2(H3K9me2/H3) or the average of H1.1AP-IC and H1.2AP-IC (H1) using STATA. M-spline divides the X-axis of a scatter plot into $10*log_{10}$(# of windows) bands and calculates the median of each band. The median points are then interpolated to fit a cubic spline. The number of bands was calculated by STATA with the total number of windows with (H3K9me2/H3) bigger than -3 and smaller than 2 or H1 bigger than -8 and smaller than 8.

4.7. Figures and tables

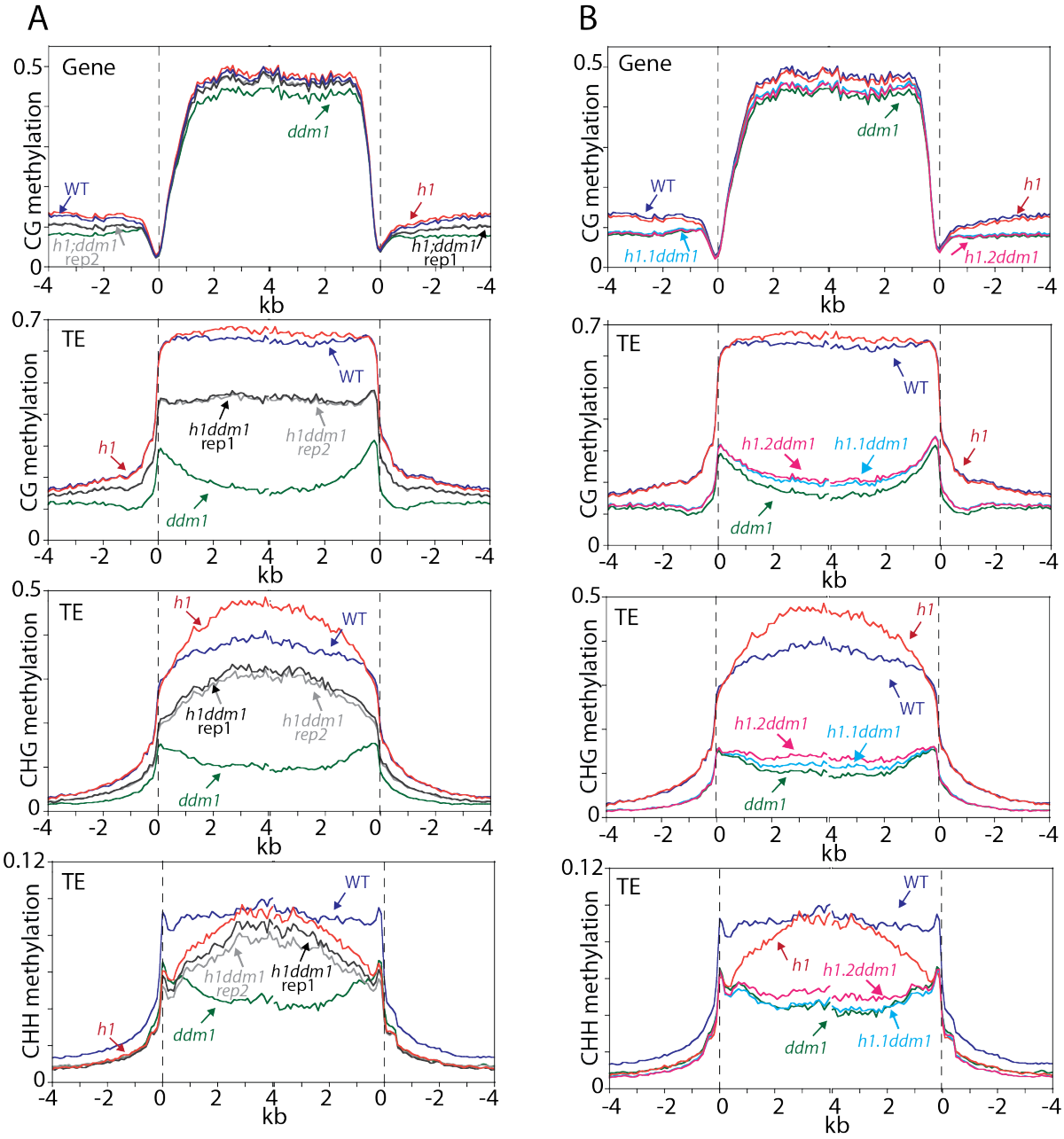

Figure 4.1. Lack of H1 reduces the methylation loss in *ddm1* plants. (A-B) Patterns of gene and transposon DNA methylation (CG, CHG, and CHH). Transposons (TE) or genes (Gene) were aligned at the 5' or the 3' end and average methylation for all cytosines within 100 bp interval is plotted. The dashed lines represent the points of alignment. Average methylation of genes and transposon in wild-type (WT), *h1*, *ddm1*, and *h1ddm1* (A) and wild-type (WT), *h1*, *ddm1*, *h1.1ddm1*, and *h1.2ddm1* (B).

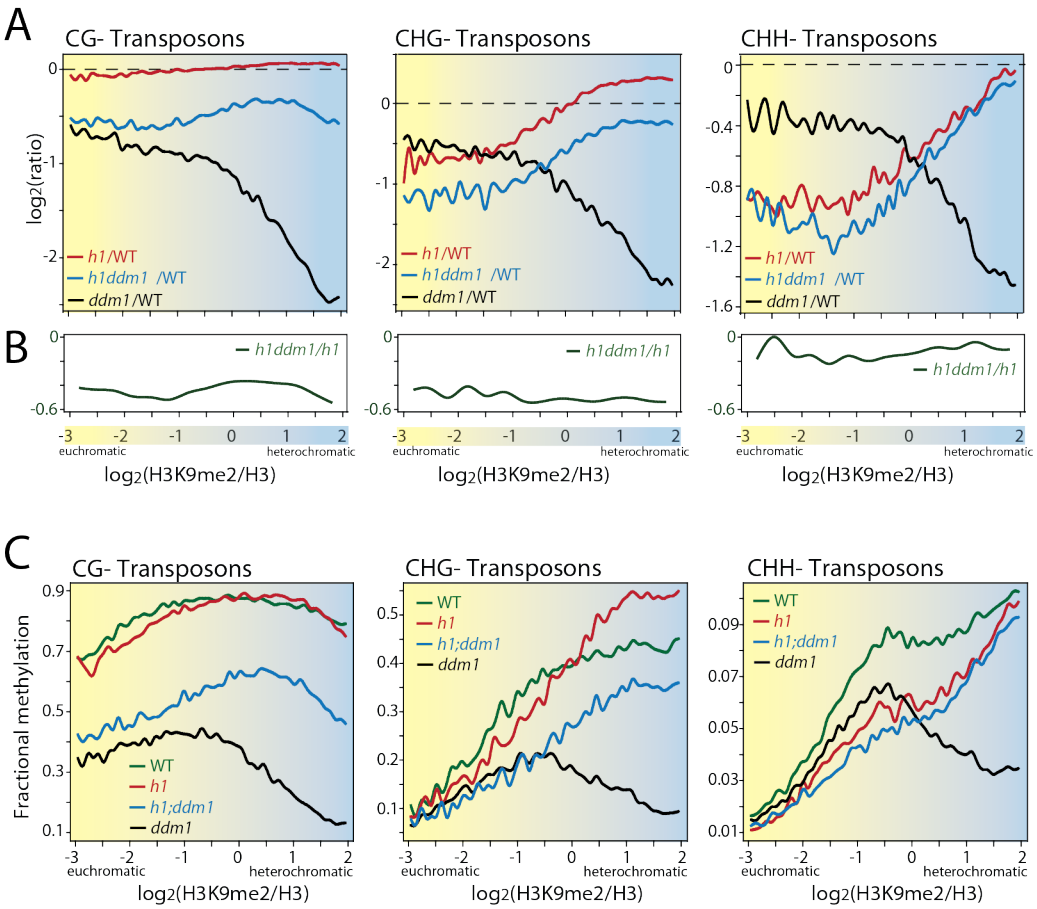

Figure 4.2. H3K9me2 enriched transposons are preferentially hypomethylated in *ddm1* but not in *h1ddm1*. (A-C) Curve fits of $\log_2$ DNA methylation ratios (A) and absolute DNA methylation levels (B) in 50 bp windows plotted against H3K9me2 level. The background gradient represents the transition from euchromatin (low H3K9me2; yellow) to heterochromatin (high H3K9me2; blue). Note that patterns of *h1ddm1* resemble that of *h1*, not *ddm1*, with the most prominent representation in H3K9me2 rich heterochromatin.

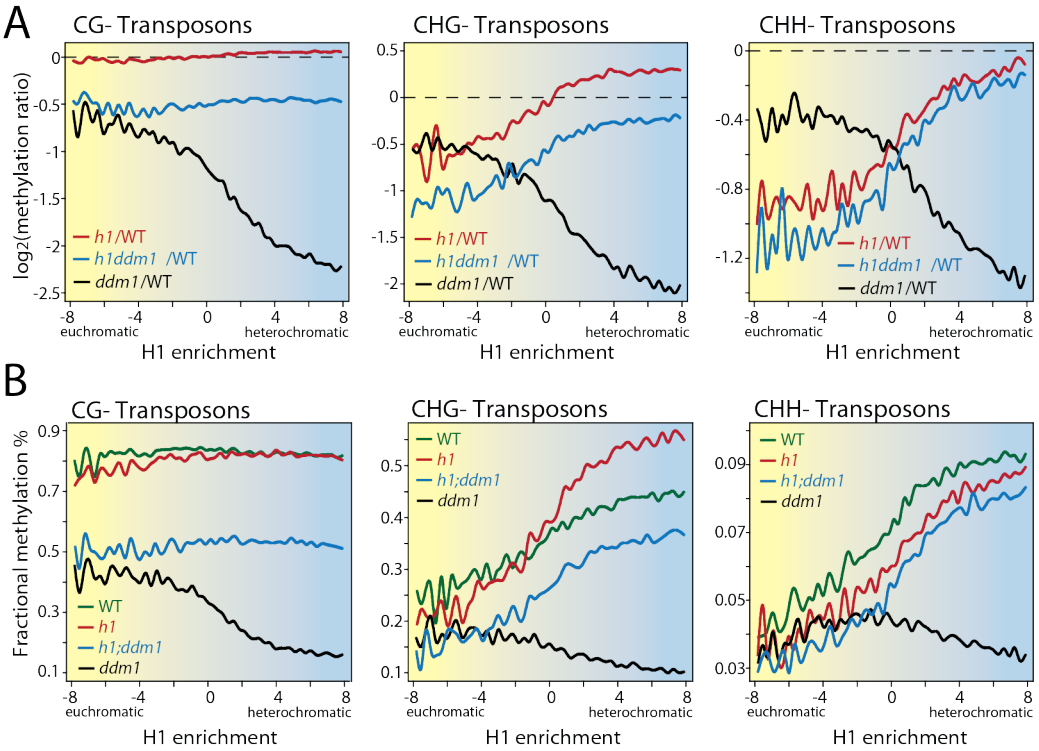

Figure 4.3. H1 enriched transposons are preferentially hypomethylated in *ddm1* but not in *h1ddm1*. (A-C) Curve fits of $\log_2$ DNA methylation ratios (A) and absolute DNA methylation levels (B) in 50 bp windows plotted against H1 level. The background gradient represents the transition from euchromatin (low H3K9me2; yellow) to heterochromatin (high H3K9me2; blue).

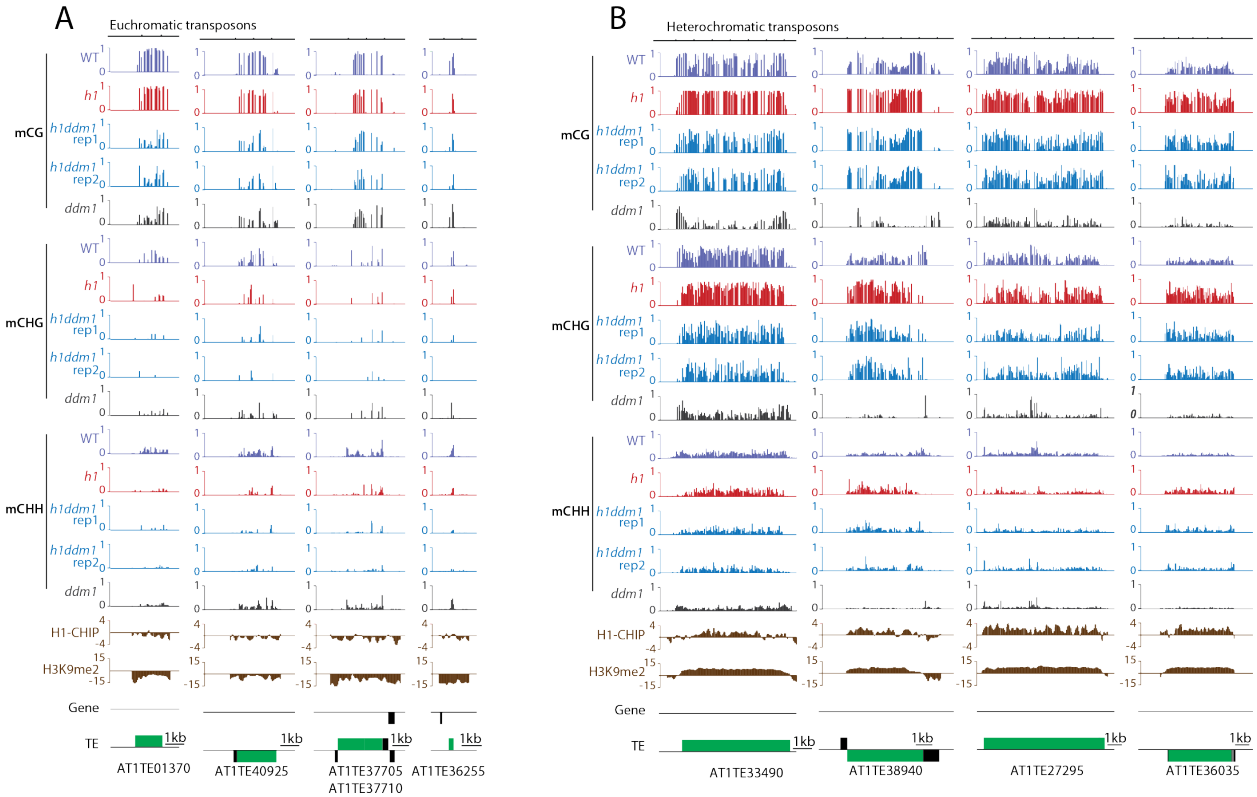

Figure 4.4. Examples of euchromatic transposons and heterochromatic transposons. (A-B) DNA methylation (CG, CHG, and CHH), H1, and H3K9me2 levels in wild-type, *h1*, *h1ddm1*-replicate1, *h1ddm1*-replicate2, and *ddm1* of euchromatic (A) and heterochromatic (B) transposons are shown.

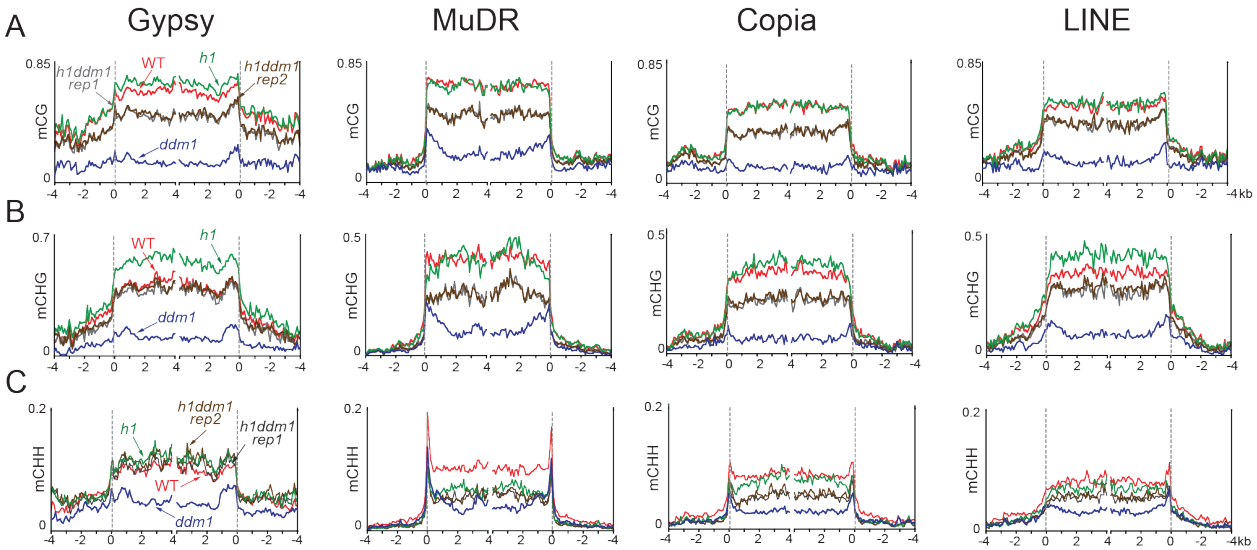

Figure 4.5. Methylation patterns of *h1ddm1* in different transposon families. (A-C) Averaged DNA methylation levels in CG (A), CHG (B), and CHH (C) are plotted as in Figure 4.1. for transposon belonging to four distinct families.

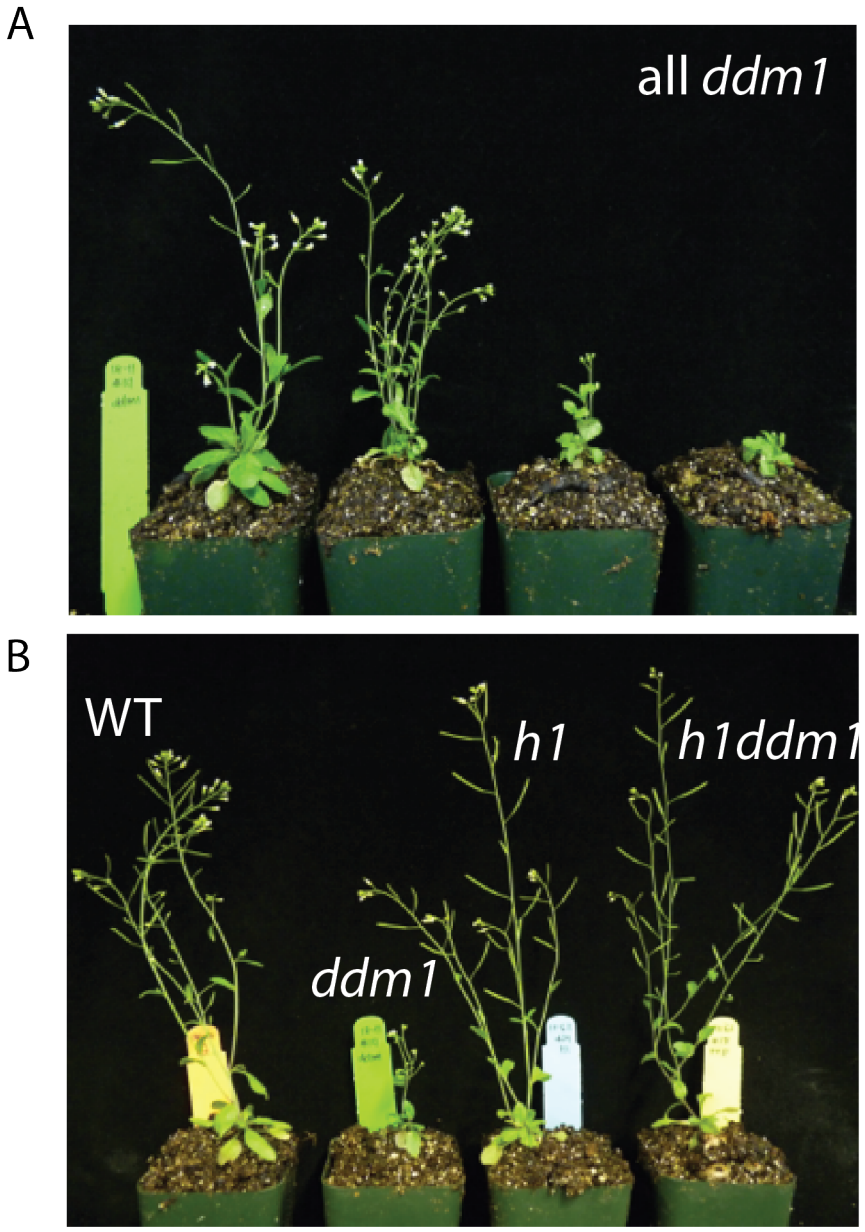

Figure 4.6. Lack of H1 in *ddm1* suppresses *ddm1* mutant phenotypes. (A) Pleiotropic phenotypic defects of *ddm1* from the cross in F3 generation. (B) Phenotypes of F3 wild-type (WT), *ddm1*, *h1*, and *h1ddm1* from the cross.

Concluding remarks

The research described in this dissertation has enhanced our knowledge on chromatin-mediated regulation of nuclear processes such as DNA demethylation, DNA methylation, and transcription. However, more questions were raised than answered in the study. Namely, is the seed DNA methylation pattern established before fertilization or after fertilization? How exactly does DDM1 influence H1 stability? What are the components that influence DNA accessibility other than H1? How does chromatin structure specify the binding of nuclear proteins? This never-ending list of questions guarantees the exciting future of follow-up studies. I believe that my dissertation work placed another stepping stone along the path towards understanding chromatin biology as a whole. It is my hope that the readers of this dissertation find the field of chromatin study as intriguing and captivating as I have.

References

Akey, C.W., and Luger, K. (2003). Histone chaperones and nucleosome assembly. Curr. Opin. Struct. Biol. *13*, 6–14.

Ammar, R., Torti, D., Tsui, K., Gebbia, M., Durbic, T., Bader, G.D., Giaever, G., and Nislow, C. (2012). Chromatin is an ancient innovation conserved between Archaea and Eukarya. Elife *1*, e00078.

Anderson, J.D., and Widom, J. (2001). Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites. Mol. Cell. Biol. *21*, 3830–3839.

Arents, G., and Moudrianakis, E.N. (1993). Topography of the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. Proc. Natl. Acad. Sci. USA *90*, 10489–10493.

Arya, G., and Schlick, T. (2009). A tale of tails: how histone tails mediate chromatin compaction in different salt and linker histone environments. J Phys Chem A *113*, 4045–4059.

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

Ascenzi, R., and GANTT, J.S. (1997). A drought-stress-inducible histone gene in *Arabidopsis thaliana* is a member of a distinct class of plant linker histone variants. Plant Mol. Biol. *34*, 629–641.

Ascenzi, R., and GANTT, J.S. (1999). Molecular genetic analysis of the drought-inducible linker histone variant in *Arabidopsis thaliana*. Plant Mol. Biol. *41*, 159–169.

Ausin, I., Mockler, T.C., Chory, J., and Jacobsen, S.E. (2009). IDN1 and IDN2 are required for de novo DNA methylation in *Arabidopsis thaliana*. Nat Struct Mol Biol *16*, 1325–1327.

Bai, L., Charvin, G., Siggia, E.D., and Cross, F.R. (2010). Nucleosome-depleted regions in cell-cycle-regulated promoters ensure reliable gene expression in every cell cycle. Dev. Cell *18*, 544–555.

Ball, D.J., Gross, D.S., and Garrard, W.T. (1983). 5-methylcytosine is localized in nucleosomes that contain histone H1. Proc. Natl. Acad. Sci. USA *80*, 5490–5494.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire,

R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120–124.

Barra, J.L., Rhounim, L., Rossignol, J.-L., and Faugeron, G. (2000). Histone H1 is dispensable for methylation-associated gene silencing in *Ascobolus immersus* and essential for long life span. Mol. Cell. Biol. *20*, 61–69.

Bates, D.L., and Thomas, J.O. (1981). Histories H1 and H5: one or two molecules per nucleosome? Nucleic Acids Res. *9*, 5883–5894.

Belikov, S., Astrand, C., and Wrange, O. (2007). Mechanism of histone H1-stimulated glucocorticoid receptor DNA binding *in vivo*. Mol. Cell. Biol. *27*, 2398–2410.

Bell, O., Tiwari, V.K., Thomä, N.H., and Schübeler, D. (2011). Determinants and dynamics of genome accessibility. Nat. Rev. Genet. *12*, 554–564.

Bellizzi, D., D'Aquila, P., Montesanto, A., Corsonello, A., Mari, V., Mazzei, B., Lattanzio, F., and Passarino, G. (2012). Global DNA methylation in old subjects is correlated with frailty. Age (Dordr) *34*, 169–179.

Bowman, G.D. (2010). Mechanisms of ATP-dependent nucleosome sliding. Curr. Opin. Struct. Biol. *20*, 73–81.

Bresnick, E.H., Bustin, M., Marsaud, V., Richard-Foy, H., and Hager, G.L. (1992). The transcriptionally-active MMTV promoter is depleted of histone H1. Nucleic Acids Res. *20*, 273–278.

Brzeski, J., and Jerzmanowski, A. (2003). Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. J. Biol. Chem. *278*, 823–828.

Buisine, N., Quesneville, H., and Colot, V. (2008). Improved detection and annotation of transposable elements in sequenced genomes using multiple reference sequence sets. Genomics *91*, 467–475.

Bureau, T.E., and Wessler, S.R. (1994). Mobile inverted-repeat elements of the Tourist family are associated with the genes of many cereal grasses. Proc. Natl. Acad. Sci. USA *91*, 1411–1415.

Burgess, R.J., and Zhang, Z. (2013). Histone chaperones in nucleosome assembly and human disease. Nat Struct Mol Biol *20*, 14–22.

Calarco, J.P., Borges, F., Donoghue, M.T.A., Van Ex, F., Jullien, P.E., Lopes, T., Gardner, R., Berger, F., Feijó, J.A., Becker, J.D., et al. (2012). Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell *151*, 194–205.

Calo, E., and Wysocka, J. (2013). Modification of enhancer chromatin: what, how, and why? Mol. Cell *49*, 825–837.

Campoy, F.J., Meehan, R.R., McKay, S., Nixon, J., and Bird, A. (1995). Binding of histone H1 to DNA is indifferent to methylation at CpG sequences. J. Biol. Chem. *270*, 26473–26481.

Chan, S.W.-L., Henderson, I.R., Zhang, X., Shah, G., Chien, J.S.-C., and Jacobsen, S.E. (2006). RNAi, DRD1, and histone methylation actively target developmentally important non-CG DNA methylation in *Arabidopsis*. PLoS Genet. *2*, e83.

Chaudhury, A.M., Koltunow, A., Payne, T., Luo, M., Tucker, M.R., Dennis, E.S., and Peacock, W.J. (2001). Control of early seed development. Annu. Rev. Cell Dev. Biol. *17*, 677–699.

Cheung, E., Zarifyan, A.S., and Kraus, W.L. (2002). Histone H1 represses estrogen receptor alpha transcriptional activity by selectively inhibiting receptor-mediated transcription initiation. Mol. Cell. Biol. *22*, 2463–2471.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer, R.L. (2002). DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell *110*, 33–42.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu. Rev. Biochem. *78*, 273–304.

Clausell, J., Happel, N., Hale, T.K., Doenecke, D., and Beato, M. (2009). Histone H1 subtypes differentially modulate chromatin condensation without preventing ATP-dependent remodeling by SWI/SNF or NURF. PLoS ONE *4*, e0007243.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature *452*, 215–219.

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

Corona, D.F.V., Siriaco, G., Armstrong, J.A., Snarskaya, N., McClymont, S.A., Scott, M.P., and Tamkun, J.W. (2007). ISWI regulates higher-order chromatin structure and histone H1 assembly *in vivo*. PLoS Biol. *5*, e232.

Danilevskaya, O.N., Hermon, P., Hantke, S., Muszynski, M.G., Kollipara, K., and Ananiev, E.V. (2003). Duplicated fie genes in maize: expression pattern and imprinting suggest distinct functions. Plant Cell *15*, 425–438.

Daujat, S., Zeissler, U., Waldmann, T., Happel, N., and Schneider, R. (2005). HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. J. Biol. Chem. *280*, 38090–38095.

de Boer, E., Rodriguez, P., Bonte, E., Krijgsveld, J., Katsantoni, E., Heck, A., Grosveld, F., and Strouboulis, J. (2003). Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. Proc. Natl. Acad. Sci. USA *100*, 7480–7485.

Deleris, A., Greenberg, M.V.C., Ausin, I., Law, R.W.Y., Moissiard, G., Schubert, D., and Jacobsen, S.E. (2010). Involvement of a Jumonji-C domain-containing histone demethylase in DRM2-mediated maintenance of DNA methylation. EMBO Rep. *11*, 950–955.

Dou, Y., and Gorovsky, M.A. (2000). Phosphorylation of linker histone H1 regulates gene expression in vivo by creating a charge patch. Mol. Cell *6*, 225–231.

Du, J., Zhong, X., Bernatavichute, Y.V., Stroud, H., Feng, S., Caro, E., Vashisht, A.A., Terragni, J., Chin, H.G., Tu, A., et al. (2012). Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell *151*, 167–180.

Dürr, H., Flaus, A., Owen-Hughes, T., and Hopfner, K.-P. (2006). Snf2 family ATPases and DExx box helicases: differences and unifying concepts from high-resolution crystal structures. Nucleic Acids Res. *34*, 4160–4167.

Eaton, M.L., Galani, K., Kang, S., Bell, S.P., and MacAlpine, D.M. (2010). Conserved nucleosome positioning defines replication origins. Genes Dev. *24*, 748–753.

Eaton, M.L., Prinz, J.A., MacAlpine, H.K., Tretyakov, G., Kharchenko, P.V., and MacAlpine, D.M. (2011). Chromatin signatures of the *Drosophila* replication program. Genome Res. *21*, 164–174.

Ebbs, M.L., and Bender, J. (2006). Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. Plant Cell *18*, 1166–1176.

Ebbs, M.L., Bartee, L., and Bender, J. (2005). H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. Mol. Cell. Biol. *25*, 10507–10515.

El-Shami, M., Pontier, D., Lahmy, S., Braun, L., Picart, C., Vega, D., Hakimi, M.-A., Jacobsen, S.E., Cooke, R., and Lagrange, T. (2007). Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev. *21*, 2539–2544.

Elmayan, T., Proux, F., and Vaucheret, H. (2005). *Arabidopsis RPA2*: a genetic link among transcriptional gene silencing, DNA repair, and DNA replication. Curr. Biol. *15*, 1919–1925.

Fan, L., and Roberts, V.A. (2006). Complex of linker histone H5 with the nucleosome and its implications for chromatin packing. Proc. Natl. Acad. Sci. USA *103*, 8384–8389.

Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L., and Skoultchi, A.I. (2005). Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell *123*, 1199–1212.

Furuyama, T., Dalal, Y., and Henikoff, S. (2006). Chaperone-mediated assembly of centromeric chromatin *in vitro*. Proc. Natl. Acad. Sci. USA *103*, 6172–6177.

Gantt, J.S., and Lenvik, T.R. (1991). *Arabidopsis thaliana* H1 histones. European Journal of Biochemistry *202*, 1029–1039.

Garcia, B.A., Busby, S.A., Barber, C.M., Shabanowitz, J., Allis, C.D., and Hunt, D.F. (2004). Characterization of phosphorylation sites on histone H1 isoforms by tandem mass spectrometry. J. Proteome Res. *3*, 1219–1227.

Gehring, M., Bubb, K.L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science *324*, 1447–1451.

Gehring, M., Huh, J.H., Hsieh, T.-F., Penterman, J., Choi, Y., Harada, J.J., Goldberg, R.B., and Fischer, R.L. (2006). DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. Cell *124*, 495–506.

Gong, Z., Morales-Ruiz, T., Ariza, R.R., Roldán-Arjona, T., David, L., and Zhu, J.-K. (2002). ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. Cell *111*, 803–814.

Görisch, S.M., Wachsmuth, M., Tóth, K.F., Lichter, P., and Rippe, K. (2005). Histone acetylation increases chromatin accessibility. J. Cell. Sci. *118*, 5825–5834.

Grewal, S.I.S., and Jia, S. (2007). Heterochromatin revisited. Nat. Rev. Genet. *8*, 35–46.

Gutiérrez-Marcos, J.F., Costa, L.M., Dal Prà, M., Scholten, S., Kranz, E., Perez, P., and Dickinson, H.G. (2006). Epigenetic asymmetry of imprinted genes in plant gametes. Nat. Genet. *38*, 876–878.

Haag, J.R., and Pikaard, C.S. (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat. Rev. Mol. Cell Biol. *12*, 483–492.

Haag, J.R., Ream, T.S., Marasco, M., Nicora, C.D., Norbeck, A.D., Pasa-Tolic, L., and Pikaard, C.S. (2012). In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. Mol. Cell *48*, 811–818.

Hale, T.K., Contreras, A., Morrison, A.J., and Herrera, R.E. (2006). Phosphorylation of the linker histone H1 by CDK regulates its binding to HP1a. Mol. Cell *22*, 693–699.

Happel, N., and Doenecke, D. (2009). Histone H1 and its isoforms: contribution to chromatin structure and function. Gene *431*, 1–12.

Haun, W.J., Laoueillé-Duprat, S., O'connell, M.J., Spillane, C., Grossniklaus, U., Phillips, A.R., Kaeppler, S.M., and Springer, N.M. (2007). Genomic imprinting, methylation and molecular evolution of maize Enhancer of zeste (Mez) homologs. Plant J. *49*, 325–337.

Hedges, S.B., Dudley, J., and Kumar, S. (2006). TimeTree: a public knowledge-base of divergence times among organisms. Bioinformatics *22*, 2971–2972.

Hellauer, K., Sirard, E., and Turcotte, B. (2001). Decreased expression of specific genes in yeast cells lacking histone H1. J. Biol. Chem. *276*, 13587–13592.

Henderson, I.R., Zhang, X., Lu, C., Johnson, L., Meyers, B.C., Green, P.J., and Jacobsen, S.E. (2006). Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat. Genet. *38*, 721–725.

Herrera, R.E., Chen, F., and Weinberg, R.A. (1996). Increased histone H1 phosphorylation and relaxed chromatin structure in Rb-deficient fibroblasts. Proc. Natl. Acad. Sci. USA *93*, 11510–11515.

Howard, P.K., Shaw, J., and Otsuka, A.J. (1985). Nucleotide sequence of the birA gene encoding the biotin operon repressor and biotin holoenzyme synthetase functions of Escherichia coli. Gene *35*, 321–331.

Hsieh, T.-F., Ibarra, C.A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R.L., and Zilberman, D. (2009). Genome-wide demethylation of *Arabidopsis* endosperm. Science *324*, 1451–1454.

Huh, J.H., Bauer, M.J., Hsieh, T.-F., and Fischer, R.L. (2008). Cellular programming of plant gene imprinting. Cell *132*, 735–744.

Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.-F., Uzawa, R., Rodrigues, J.A., Zemach, A., Chumak, N., Machlicova, A., Nishimura, T., et al. (2012). Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. Science *337*, 1360–1364.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556–560.

Jacob, Y., Stroud, H., Leblanc, C., Feng, S., Zhuo, L., Caro, E., Hassel, C., Gutierrez, C., Michaels, S.D., and Jacobsen, S.E. (2010). Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. Nature *466*, 987–991.

Jacobs, A.L., and Schär, P. (2012). DNA glycosylases: in DNA repair and beyond. Chromosoma *121*, 1–20.

Jahnke, S., and Scholten, S. (2009). Epigenetic resetting of a gene imprinted in plant embryos. Curr. Biol. *19*, 1677–1681.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999). Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat. Genet. *22*, 94–97.

Jedrusik, M.A., and Schulze, E. (2001). A single histone H1 isoform (H1.1) is essential for chromatin silencing and germline development in *Caenorhabditis elegans*. Development *128*, 1069–1080.

Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. Nat. Rev. Genet. *10*, 161–172.

Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S.R., McCouch, S.R., and Wessler, S.R. (2003). An active DNA transposon family in rice. Nature *421*, 163–167.

Johnson, L.M., Bostick, M., Zhang, X., Kraft, E., Henderson, I., Callis, J., and Jacobsen, S.E. (2007). The SRA methyl-cytosine-binding domain links DNA and histone methylation. Curr. Biol. *17*, 379–384.

Juan, L.J., Utley, R.T., Adams, C.C., Vettese-Dadey, M., and Workman, J.L. (1994). Differential repression of transcription factor binding by histone H1 is regulated by the core histone amino termini. Embo J. *13*, 6031–6040.

Juan, L.J., Utley, R.T., Vignali, M., Bohm, L., and Workman, J.L. (1997). H1-mediated repression of transcription factor binding to a stably positioned nucleosome. J. Biol. Chem. *272*, 3635–3640.

Jurkowski, T.P., and Jeltsch, A. (2011). On the evolutionary origin of eukaryotic

DNA methyltransferases and Dnmt2. PLoS ONE *6*, e28104.

Kakutani, T., Jeddeloh, J.A., Flowers, S.K., Munakata, K., and Richards, E.J. (1996). Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc. Natl. Acad. Sci. USA *93*, 12406–12411.

Kakutani, T., Munakata, K., Richards, E.J., and Hirochika, H. (1999). Meiotically and mitotically stable inheritance of DNA hypomethylation induced by ddm1 mutation of *Arabidopsis thaliana*. Genetics *151*, 831–838.

Kanno, T., Mette, M.F., Kreil, D.P., Aufsatz, W., Matzke, M., and Matzke, A.J.M. (2004). Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. Curr. Biol. *14*, 801–805.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., et al. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. Nature *458*, 362–366.

Kasinsky, H.E., Lewis, J.D., Dacks, J.B., and Ausio, J. (2001). Origin of H1 linker histones. FASEB J. *15*, 34–42.

Khalequzzaman, M., and Haq, N. (2005). Isolation and in vitro fusion of egg and sperm cells in Oryza sativa. Plant Physiol. Biochem. *43*, 69–75.

Knizewski, L., Ginalski, K., and Jerzmanowski, A. (2008). Snf2 proteins in plants: gene silencing and beyond. Trends in Plant Science *13*, 557–565.

Kodama, Y., Nagaya, S., Shinmyo, A., and Kato, K. (2007). Mapping and characterization of DNase I hypersensitive sites in *Arabidopsis* chromatin. Plant Cell Physiol. *48*, 459–470.

Kornberg, R.D., and Lorch, Y. (1992). Chromatin structure and transcription. Annu. Rev. Cell Biol. *8*, 563–587.

Kranz, E., Bautor, J., and Lörz, H. (1991). In vitro fertilization of single, isolated gametes of maize mediated by electrofusion. Sex Plant Reprod *4*, 12–16.

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.

la Paz Sanchez, de, M., and Gutierrez, C. (2009). *Arabidopsis* ORC1 is a PHD-containing H3K4me3 effector that regulates transcription. Proc. Natl. Acad. Sci. USA *106*, 2065–2070.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. *11*, 204–220.

Law, J.A., Ausin, I., Johnson, L.M., Vashisht, A.A., Zhu, J.-K., Wohlschlegel, J.A., and Jacobsen, S.E. (2010). A Protein Complex Required for Polymerase V Transcripts and RNA-Directed DNA Methylation in *Arabidopsis*. Curr. Biol. *20*, 951–956.

Law, J.A., Vashisht, A.A., Wohlschlegel, J.A., and Jacobsen, S.E. (2011). SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. PLoS Genet. *7*, e1002195.

Laybourn, P.J., and Kadonaga, J.T. (1991). Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. Science *254*, 238–245.

Lee, W.Y., Lee, D., Chung, W.-I., and Kwon, C.S. (2009). *Arabidopsis* ING and Alfin1-like protein families localize to the nucleus and bind to H3K4me3/2 via plant homeodomain fingers. Plant J. *58*, 511–524.

Lever, M.A., Th'ng, J.P., Sun, X., and Hendzel, M.J. (2000). Rapid exchange of histone H1.1 on chromatin in living human cells. Nature *408*, 873–876.

Li, B., Carey, M., and Workman, J.L. (2007a). The role of chromatin during transcription. Cell *128*, 707–719.

Li, C.F., Henderson, I.R., Song, L., Fedoroff, N., Lagrange, T., and Jacobsen, S.E. (2008). Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in *Arabidopsis thaliana*. PLoS Genet. *4*, e27.

Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y.V., Chan, S.W.-L., Lagrange, T., Pikaard, C.S., and Jacobsen, S.E. (2006). An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. Cell *126*, 93–106.

Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Curr. Opin. Genet. Dev. *21*, 175–186.

Li, H., Freeling, M., and Lisch, D. (2010). Epigenetic reprogramming during vegetative phase change in maize. Proc. Natl. Acad. Sci. USA *107*, 22184–22189.

Li, J., Hsia, A.-P., and Schnable, P.S. (2007b). Recent advances in plant recombination. Curr. Opin. Plant Biol. *10*, 131–135.

Li, J.-Y., Patterson, M., Mikkola, H.K.A., Lowry, W.E., and Kurdistani, S.K. (2012). Dynamic distribution of linker histone H1.5 in cellular differentiation. PLoS Genet. *8*, e1002879.

Lippman, Z., Gendrel, A.-V., Black, M., Vaughn, M.W., Dedhia, N., McCombie,

W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D., et al. (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature *430*, 471–476.

Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., and Ecker, J.R. (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. Cell *133*, 523–536.

Liu, C., Lu, F., Cui, X., and Cao, X. (2010). Histone methylation in higher plants. Annu. Rev. Plant Biol. *61*, 395–420.

Liu, X., Gao, Q., Li, P., Zhao, Q., Zhang, J., Li, J., Koseki, H., and Wong, J. (2013). UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. Nat. Comms. *4*, 1563.

Loidl, P. (2004). A plant dialect of the histone language. Trends in Plant Science *9*, 84–90.

Losada, A., and Hirano, T. (2005). Dynamic molecular linkers of the genome: the first decade of SMC proteins. Genes Dev. *19*, 1269–1287.

Lu, X., Wontakal, S.N., Emelyanov, A.V., Morcillo, P., Konev, A.Y., Fyodorov, D.V., and Skoultchi, A.I. (2009). Linker histone H1 is essential for *Drosophila* development, the establishment of pericentric heterochromatin, and a normal polytene chromosome structure. Genes Dev. *23*, 452–465.

Lu, X., Wontakal, S.N., Kavi, H., Kim, B.J., Guzzardo, P.M., Emelyanov, A.V., Xu, N., Hannon, G.J., Zavadil, J., Fyodorov, D.V., et al. (2013). *Drosophila* H1 regulates the genetic activity of heterochromatin by recruitment of Su(var)3-9. Science *340*, 78–81.

Lu, X., and Hansen, J.C. (2004). Identification of specific functional subdomains within the linker histone H10 C-terminal domain. J. Biol. Chem. *279*, 8701–8707.

Luger, K., Mäder, 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.

Luk, E., Ranjan, A., Fitzgerald, P.C., Mizuguchi, G., Huang, Y., Wei, D., and Wu, C. (2010). Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. Cell *143*, 725–736.

Maegawa, S., Hinkal, G., Kim, H.S., Shen, L., Zhang, L., Zhang, J., Zhang, N., Liang, S., Donehower, L.A., and Issa, J.-P.J. (2010). Widespread and tissue specific age-related DNA methylation changes in mice. Genome Res. *20*, 332–340.

Malagnac, F., Bartee, L., and Bender, J. (2002). An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. EMBO J. *21*, 6842–6852.

Margueron, R., Trojer, P., and Reinberg, D. (2005). The key to development: interpreting the histone code? Curr. Opin. Genet. Dev. *15*, 163–176.

Martienssen, R., Barkan, A., Taylor, W.C., and Freeling, M. (1990). Somatically heritable switches in the DNA modification of Mu transposable elements monitored with a suppressible mutant in maize. Genes Dev. *4*, 331–343.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat. Rev. Mol. Cell Biol. *6*, 838–849.

Mathieu, O., and Bender, J. (2004). RNA-directed DNA methylation. J. Cell. Sci. *117*, 4881–4888.

Mathieu, O., Probst, A.V., and Paszkowski, J. (2005). Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis*. Embo J. *24*, 2783–2791.

McArthur, M., and Thomas, J.O. (1996). A preference of histone H1 for methylated DNA. EMBO J. *15*, 1705–1714.

McGhee, J.D., and Felsenfeld, G. (1980). Nucleosome structure. Annu. Rev. Biochem. *49*, 1115–1156.

Meyer, P., Linn, F., Heidmann, I., Meyer, H., Niedenhof, I., and Saedler, H. (1992). Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. Mol. Gen. Genet. *231*, 345–352.

Misteli, T., Gunjan, A., Hock, R., Bustin, M., and Brown, D.T. (2000). Dynamic binding of histone H1 to chromatin in living cells. Nature *408*, 877–881.

Mito, Y., Henikoff, J.G., and Henikoff, S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. Nat. Genet. *37*, 1090–1097.

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 *336*, 1448–1451.

Molina, C., and Grotewold, E. (2005). Genome wide analysis of *Arabidopsis* core promoters. BMC Genomics *6*, 25.

Mosher, R.A., Melnyk, C.W., Kelly, K.A., Dunn, R.M., Studholme, D.J., and Baulcombe, D.C. (2009). Uniparental expression of PolIV-dependent siRNAs in

developing endosperm of *Arabidopsis*. Nature *460*, 283–286.

Mylne, J.S., Barrett, L., Tessadori, F., Mesnage, S., Johnson, L., Bernatavichute, Y.V., Jacobsen, S.E., Fransz, P., and Dean, C. (2006). LHP1, the *Arabidopsis* homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. Proc. Natl. Acad. Sci. USA *103*, 5012–5017.

Ni, J.-Q., Liu, L.-P., Hess, D., Rietdorf, J., and Sun, F.-L. (2006). *Drosophila* ribosomal proteins are associated with linker histone H1 and suppress gene transcription. Genes Dev. *20*, 1959–1973.

Nielsen, A.L., Oulad-Abdelghani, M., Ortiz, J.A., Remboutsika, E., Chambon, P., and Losson, R. (2001). Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. Mol. Cell *7*, 729–739.

Nightingale, K.P., Wellinger, R.E., Sogo, J.M., and Becker, P.B. (1998). Histone acetylation facilitates RNA polymerase II transcription of the *Drosophila* hsp26 gene in chromatin. Embo J. *17*, 2865–2876.

O'Neill, T.E., Meersseman, G., Pennings, S., and Bradbury, E.M. (1995). Deposition of histone H1 onto reconstituted nucleosome arrays inhibits both initiation and elongation of transcripts by T7 RNA polymerase. Nucleic Acids Res. *23*, 1075–1082.

Olins, A.L., and Olins, D.E. (1974). Spheroid Chromatin Units (ngr Bodies). Science *183*, 330–332.

Ono, A., Yamaguchi, K., Fukada-Tanaka, S., Terada, R., Mitsui, T., and Iida, S. (2012). A null mutation of ROS1a for DNA demethylation in rice is not transmittable to progeny. Plant J. *71*, 564–574.

Papamichos-Chronakis, M., Watanabe, S., Rando, O.J., and Peterson, C.L. (2011). Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. Cell *144*, 200–213.

Pennings, S., Meersseman, G., and Bradbury, E.M. (1994). Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. Proc. Natl. Acad. Sci. USA *91*, 10275–10279.

Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L. (2007). DNA demethylation in the *Arabidopsis* genome. Proc. Natl. Acad. Sci. USA *104*, 6752–6757.

Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. Nature *433*, 434–438.

Probst, A.V., Fagard, M., Proux, F., Mourrain, P., Boutet, S., Earley, K., Lawrence, R.J., Pikaard, C.S., Murfett, J., Furner, I., et al. (2004). *Arabidopsis* histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. Plant Cell *16*, 1021–1034.

Qian, W., Miki, D., Zhang, H., Liu, Y., Zhang, X., Tang, K., Kan, Y., La, H., Li, X., Li, S., et al. (2012). A histone acetyltransferase regulates active DNA demethylation in *Arabidopsis*. Science *336*, 1445–1448.

Qu, L.Q., Xing, Y.P., Liu, W.X., Xu, X.P., and Song, Y.R. (2008). Expression pattern and activity of six glutelin gene promoters in transgenic rice. J. Exp. Bot. *59*, 2417–2424.

Raghuram, N., Carrero, G., Th'ng, J., and Hendzel, M.J. (2009). Molecular dynamics of histone H1. Biochem. Cell Biol. *87*, 189–206.

Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L., and Sweet, R.M. (1993). Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. Nature *362*, 219–223.

Rea, M., Zheng, W., Chen, M., Braud, C., Bhangu, D., Rognan, T.N., and Xiao, W. (2012). Histone H1 affects gene imprinting and DNA methylation in *Arabidopsis*. Plant J. *71*, 776–786.

Ream, T.S., Haag, J.R., Wierzbicki, A.T., Nicora, C.D., Norbeck, A.D., Zhu, J.-K., Hagen, G., Guilfoyle, T.J., Pasa-Tolic, L., and Pikaard, C.S. (2009). Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. Mol. Cell *33*, 192–203.

Robinson, P.J.J., Fairall, L., Huynh, V.A.T., and Rhodes, D. (2006). EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. Proc. Natl. Acad. Sci. USA *103*, 6506–6511.

Rodrigues, J.A., Ruan, R., Nishimura, T., Sharma, M.K., Sharma, R., Ronald, P.C., Fischer, R.L., and Zilberman, D. (2013). Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. Proc. Natl. Acad. Sci. USA *110*, 7934–7939.

Rosa, M., Harder, Von, M., Cigliano, R.A., Schlögelhofer, P., and Scheid, O.M. (2013). The *Arabidopsis* SWR1 Chromatin-Remodeling Complex Is Important for DNA Repair, Somatic Recombination, and Meiosis. Plant Cell *25*, 1990-2001.

Routh, A., Sandin, S., and Rhodes, D. (2008). Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proc. Natl. Acad. Sci. USA *105*, 8872–8877.

Rupp, R.A.W., and Becker, P.B. (2005). Gene regulation by histone H1: new links to DNA methylation. Cell *123*, 1178–1179.

Ryan, D.P., and Owen-Hughes, T. (2011). Snf2-family proteins: chromatin remodellers for any occasion. Curr. Opin. Chem. Biol. *15*, 649–656.

Schoft, V.K., Chumak, N., Mosiolek, M., Slusarz, L., Komnenovic, V., Brownfield, L., Twell, D., Kakutani, T., and Tamaru, H. (2009). Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. EMBO Rep. *10*, 1015–1021.

Schones, D.E., and Zhao, K. (2008). Genome-wide approaches to studying chromatin modifications. Nat. Rev. Genet. *9*, 179–191.

Schwabish, M.A., and Struhl, K. (2007). The Swi/Snf complex is important for histone eviction during transcriptional activation and RNA polymerase II elongation in vivo. Mol. Cell. Biol. *27*, 6987–6995.

Schwaiger, M., Stadler, M.B., Bell, O., Kohler, H., Oakeley, E.J., and Schübeler, D. (2009). Chromatin state marks cell-type- and gender-specific replication of the *Drosophila* genome. Genes Dev. *23*, 589–601.

Schwartz, S., Meshorer, E., and Ast, G. (2009). Chromatin organization marks exon-intron structure. Nat. Struct. Mol. Biol. *16*, 990–995.

Searle, I.R., Pontes, O., Melnyk, C.W., Smith, L.M., and Baulcombe, D.C. (2010). JMJ14, a JmjC domain protein, is required for RNA silencing and cell-to-cell movement of an RNA silencing signal in *Arabidopsis*. Genes Dev. *24*, 986–991.

Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.-P.Z., and Widom, J. (2006). A genomic code for nucleosome positioning. Nature *442*, 772–778.

Shahbazian, M.D., and Grunstein, M. (2007). Functions of site-specific histone acetylation and deacetylation. Annu. Rev. Biochem. *76*, 75–100.

Shen, X., and Gorovsky, M.A. (1996). Linker histone H1 regulates specific gene expression but not global transcription in vivo. Cell *86*, 475–483.

Shen, X., Yu, L., Weir, J.W., and Gorovsky, M.A. (1995). Linker histories are not essential and affect chromatin condensation in vivo. Cell *82*, 47–56.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzić, M., Becker, J.D., Feijó, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell *136*, 461–472.

Smith, L.M., Pontes, O., Searle, I., Yelina, N., Yousafzai, F.K., Herr, A.J., Pikaard,

C.S., and Baulcombe, D.C. (2007). An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. Plant Cell *19*, 1507–1521.

Spiker, S., Murray, M.G., and Thompson, W.F. (1983). DNase I sensitivity of transcriptionally active genes in intact nuclei and isolated chromatin of plants. Proc. Natl. Acad. Sci. USA *80*, 815–819.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. Nat. Struct. Mol. Biol. *20*, 267–273.

Suto, R.K., Edayathumangalam, R.S., White, C.L., Melander, C., Gottesfeld, J.M., Dervan, P.B., and Luger, K. (2003). Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. J. Mol. Biol. *326*, 371–380.

Swanson, W.J., and Vacquier, V.D. (2002). The rapid evolution of reproductive proteins. Nat. Rev. Genet. *3*, 137–144.

Takasuka, T.E., and Stein, A. (2010). Direct measurements of the nucleosome-forming preferences of periodic DNA motifs challenge established models. Nucleic Acids Res. *38*, 5672–5680.

Takeda, S., Tadele, Z., Hofmann, I., Probst, A.V., Angelis, K.J., Kaya, H., Araki, T., Mengiste, T., Mittelsten Scheid, O., Shibahara, K.-I., et al. (2004). BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in *Arabidopsis*. Genes Dev. *18*, 782–793.

Tao, Y., Xi, S., Shan, J., Maunakea, A., Che, A., Briones, V., Lee, E.Y., Geiman, T., Huang, J., Stephens, R., et al. (2011). Lsh, chromatin remodeling family member, modulates genome-wide cytosine methylation patterns at nonrepeat sequences. Proc. Natl. Acad. Sci. USA *108*, 5626–5631.

Thoma, F. (2005). Repair of UV lesions in nucleosomes--intrinsic properties and remodeling. DNA Repair (Amst.) *4*, 855–869.

Thomas, J.O. (1999). Histone H1: location and role. Curr. Opin. Cell Biol. *11*, 312–317.

Tolkunov, D., and Morozov, A.V. (2009). Nucleosome positioning and energetics: Recent advances in genomic and computational studies. arXiv Preprint arXiv:0912.3954.

Tran, R.K., Zilberman, D., de Bustos, C., Ditt, R.F., Henikoff, J.G., Lindroth, A.M., Delrow, J., Boyle, T., Kwong, S., Bryson, T.D., et al. (2005). Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in *Arabidopsis*. Genome Biol. *6*, R90.

Tsukahara, S., Kobayashi, A., Kawabe, A., Mathieu, O., Miura, A., and Kakutani, T. (2009). Bursts of retrotransposition reproduced in *Arabidopsis*. Nature *461*, 423–426.

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 Genet. *3*, e86.

Uchiumi, T., Komatsu, S., Koshiba, T., and Okamoto, T. (2006). Isolation of gametes and central cells from Oryza sativa L. Sex Plant Reprod. *19*, 37–45.

Ura, K., Araki, M., Saeki, H., Masutani, C., Ito, T., Iwai, S., Mizukoshi, T., Kaneda, Y., and Hanaoka, F. (2001). ATP-dependent chromatin remodeling facilitates nucleotide excision repair of UV-induced DNA lesions in synthetic dinucleosomes. EMBO J. *20*, 2004–2014.

Vaillant, I., Schubert, I., Tourmente, S., and Mathieu, O. (2006). MOM1 mediates DNA-methylation-independent silencing of repetitive sequences in *Arabidopsis*. EMBO Rep. *7*, 1273–1278.

van Attikum, H., and Gasser, S.M. (2005). The histone code at DNA breaks: a guide to repair? Nat. Rev. Mol. Cell Biol. *6*, 757–765.

Vignali, M., and Workman, J.L. (1998). Location and function of linker histones. Nat. Struct. Biol. *5*, 1025–1028.

Vincent, J.A., Kwong, T.J., and Tsukiyama, T. (2008). ATP-dependent chromatin remodeling shapes the DNA replication landscape. Nat. Struct. Mol. Biol. *15*, 477–484.

Wan, J., Lin, J., Zack, D.J., and Qian, J. (2009). Relating periodicity of nucleosome organization and gene regulation. Bioinformatics *25*, 1782–1788.

Weintraub, H. (1984). Histone-H1-dependent chromatin superstructures and the suppression of gene activity. Cell *38*, 17–27.

Whitehouse, I., Rando, O.J., Delrow, J., and Tsukiyama, T. (2007). Chromatin remodelling at promoters suppresses antisense transcription. Nature *450*, 1031–1035.

Wierzbicki, A.T., and Jerzmanowski, A. (2005). Suppression of histone H1 genes in *Arabidopsis* results in heritable developmental defects and stochastic changes in DNA methylation. Genetics *169*, 997–1008.

Wierzbicki, A.T., Haag, J.R., and Pikaard, C.S. (2008). Noncoding transcription by

RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell *135*, 635–648.

Wierzbicki, A.T., Ream, T.S., Haag, J.R., and Pikaard, C.S. (2009). RNA polymerase V transcription guides ARGONAUTE4 to chromatin. Nat. Genet. *41*, 630–634.

Winston, F., and Carlson, M. (1992). Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. Trends Genet. *8*, 387–391.

Woo, H.R., Dittmer, T.A., and Richards, E.J. (2008). Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in *Arabidopsis*. PLoS Genet. *4*, e1000156.

Woo, H.R., Pontes, O., Pikaard, C.S., and Richards, E.J. (2007). VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. Genes Dev. *21*, 267–277.

Woodcock, C.L., Safer, J.P., and Stanchfield, J.E. (1976). Structural repeating units in chromatin. I. Evidence for their general occurrence. Exp. Cell Res. *97*, 101–110.

Woodcock, C.L., and Ghosh, R.P. (2010). Chromatin higher-order structure and dynamics. Cold Spring Harb Perspect Biol *2*, a000596.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., et al. (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86–90.

Xue, L.-J., Zhang, J.-J., and Xue, H.-W. (2009). Characterization and expression profiles of miRNAs in rice seeds. Nucleic Acids Res. *37*, 916–930.

Yamanaka, S. (2012). Induced pluripotent stem cells: past, present, and future. Cell Stem Cell *10*, 678–684.

Yang, C., Bolotin, E., Jiang, T., Sladek, F.M., and Martinez, E. (2007). Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene *389*, 52–65.

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.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science *328*, 916–919.

Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by

histone modifications. Nat. Struct. Mol. Biol. *20*, 259–266.

Zhang, H., Ma, Z.-Y., Zeng, L., Tanaka, K., Zhang, C.-J., Ma, J., Bai, G., Wang, P., Zhang, S.-W., Liu, Z.-W., et al. (2013). DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. Proc. Natl. Acad. Sci. USA *110*, 8290–8295.

Zhang, H., and Zhu, J.-K. (2011). RNA-directed DNA methylation. Curr. Opin. Plant Biol. *14*, 142–147.

Zhang, W., Zhang, T., Wu, Y., and Jiang, J. (2012). Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in *Arabidopsis*. Plant Cell *24*, 2719–2731.

Zhang, X., Germann, S., Blus, B.J., Khorasanizadeh, S., Gaudin, V., and Jacobsen, S.E. (2007). The *Arabidopsis* LHP1 protein colocalizes with histone H3 Lys27 trimethylation. Nat. Struct. Mol. Biol. *14*, 869–871.

Zheng, Y., John, S., Pesavento, J.J., Schultz-Norton, J.R., Schiltz, R.L., Baek, S., Nardulli, A.M., Hager, G.L., Kelleher, N.L., and Mizzen, C.A. (2010). Histone H1 phosphorylation is associated with transcription by RNA polymerases I and II. The J. Cell Biol. *189*, 407–415.

Zhou, J., Wang, X., He, K., Charron, J.-B.F., Elling, A.A., and Deng, X.W. (2010). Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in *Arabidopsis* reveals correlation between multiple histone marks and gene expression. Plant Mol. Biol. *72*, 585–595.

Zhou, Y.B., Gerchman, S.E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. Nature *395*, 402–405.

Zhu, J.-K. (2009). Active DNA demethylation mediated by DNA glycosylases. Annu. Rev. Genet. *43*, 143–166.

Zilberman, D., Coleman-Derr, D., Ballinger, T., and Henikoff, S. (2008). Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature *456*, 125–129.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. *39*, 61–69.