

RNA-directed DNA methylation

Mechanisms and functions

Magdy M. Mahfouz

Center for Plant Stress Genomics & Technology; 4700 King Abdullah University of Science & Technology (KAUST); Thuwal, Kingdom of Saudi Arabia; and Agricultural Genetic Engineering Research Institute (AGERI); Agricultural Research Center; Giza, Egypt

Key words: epigenetics, RdDM, DNA methylation, DNA demethylation, epialleles, abiotic stress, transgenerational inheritance

Epigenetic RNA based gene silencing mechanisms play a major role in genome stability and control of gene expression. Transcriptional gene silencing via RNA-directed DNA methylation (RdDM) guides the epigenetic regulation of the genome in response to disease states, growth, developmental and stress signals. RdDM machinery is composed of proteins that produce and modify 24-nt-long siRNAs, recruit the RdDM complex to genomic targets, methylate DNA and remodel chromatin. The final DNA methylation pattern is determined by either DNA methyltransferase alone or by the combined action of DNA methyltransferases and demethylases. The dynamic interaction between RdDM and demethylases may render the plant epigenome plastic to growth, developmental and environmental cues. The epigenome plasticity may allow the plant genome to assume many epigenomes and to have the right epigenome at the right time in response to intracellular or extracellular stimuli. This review discusses recent advances in RdDM research and considers future perspectives.

Most of the DNA sequences in the mammalian and plant genomes are transcribed into non-coding RNAs (ncRNAs). The recently discovered roles of the ncRNAs in gene regulation processes have revolutionized our understanding of molecular, cellular, developmental and evolutionary biology. Most of these ncRNAs are of different sizes, sequences, genomic loci and biogenesis and they all facilitate sequence specific gene silencing. An expanded and diverse ncRNAs-based gene silencing machineries operate in plants. The ncRNAs and their silencing machinery may function in a cell or tissue specific manner, at certain developmental stage and in response to a disease state or adverse environmental cues. Four RNA silencing pathways have been characterized in Arabidopsis namely microRNA (miRNA), transacting short interfering RNA (ta-siRNA), natural antisense transcript derived siRNA (nat-siRNA) and heterochromatic interfering RNA (siRNA). Unlike miRNA, ta-siRNA and nat-siRNA which function at the post-transcriptional level through mRNA degradation or translational repression, heterochromatic siRNA

mediates the gene silencing at transcriptional level by directing DNA methylation.

dsRNAs can be produced from overlapping and inverted repeats transcription by RNA polymerase II (RNA Pol II). dsRNA formation can also be initiated by the plant specific RNA Pol IV. These dsRNAs are then processed by the gene silencing machinery to generate short interfering RNAs (siRNAs) which guide DNA and chromatin modifications on homologous sequences in the genome. RNA-directed DNA methylation (RdDM) pathway which may be unique to plants cells, similar and related process takes place in mammals, functions to mediate epigenetic modifications. The RdDM process was first discovered in plants infected with viroids.¹ Plants use methylation epigenetic modifications to control gene expression possibly as an adaptation mechanism to ensure survival under unfavorable conditions. Methylation dependent silencing at transposons sequences protects the genome. Methylation in gene body regulatory sequences regulate gene expression and ultimately various cellular processes important to growth, development and survival. The epigenetic methylation marks occur at the cytosine nucleotide of the DNA and may be inherited over many generations. Three cytosine methylation systems operate in the Arabidopsis genome in three different sequence contexts: CpG, CpHpG and CpHpH (H is adenine A, thymine T or cytosine C). The methylation can be directed by siRNAs to cause gene silencing. The methylation at some target loci can be reversed by the action of cytosine demethylases. The methylation/demethylation systems provide a dynamic control of gene expression patterns and hence genome plasticity in response to various developmental, growth and stress signals. Throughout this review we will highlight recent advances in RdDM mechanisms and functions pertinent to development, stress tolerance and genome defense.

The Key Players of the RdDM Silencing Machinery

DNA methylation of cytosine residues is required by all multicellular organisms to maintain normal development and proper responses to disease states and environmental cues. Targeted disruption of various DNA methyltransferases result in developmental abnormalities.²⁻⁴ DNA methylation is known to be the major modification of the plant genome but information on how the methylation marks is directed to certain parts of the genome is unclear.^{5,6} RdDM provided the first evidence that RNA molecules can feedback on the genome and induce epigenetic modification

Correspondence to: Magdy M. Mahfouz; Email: magdy.mahfouz@kaust.edu.sa

Submitted: 03/02/10; Accepted: 03/03/10

Previously published online:

www.landesbioscience.com/journals/psb/article/11695

to their cognate DNA sequences. In the cell nucleus, siRNAs of sequences homologous to DNA in the genome are used to guide a specific and often reversible DNA methylation.

The RdDM methylation machinery is composed of several proteins that process and produce the siRNAs, modify the histones, remodel the chromatin and methylate the cytosine in all sequence contexts. The siRNAs are produced from dsRNAs. The dsRNA can be generated from inverted repeats or overlapping transcription by RNA Pol III. Pol IV transcription of methylated DNA can produce single stranded RNA (ssRNA) which is converted into double stranded RNA (dsRNA) by the RNA dependent RNA polymerase 2 (RDR2). The dsRNAs are then processed by DCL3 followed by the HEN1 methylation and loaded on AGO4 which interacts with WG/GW CTD of the largest subunit of Pol V, NRPE1.⁷ The siRNAs guided AGO4 and its associated proteins to the homologous DNA sequences and facilitate target methylation by the domains rearranged 2 (DRM2) de novo methyltransferase. AGO4/siRNA may interact with the nascent Pol V transcripts or DNA to guide the methylation process.⁸ The Pol IV generates precursor transcripts that produce and or amplify the siRNA trigger. In conclusion, The RdDM machinery is composed of DNA methyl transferases, histone modifying enzymes, chromatin remodelling proteins, plant specific RNA polymerases (Pol IV and Pol V) and RNAi machinery proteins (RDR2, DCL3, HEN1 and AGO4/6). I will address in more detail the molecular structure and function of these key players of the RdDM machinery in the following sections.

Plant specific Pol IV and Pol V polymerases. Eukaryotic organisms have three multisubunit DNA-dependent RNA polymerases with distinct functions namely RNA Pol I, II and III. RNA Pol I transcribes the 45S ribosomal DNA genes to generate 45S ribosomal RNAs, RNA Pol II transcribes the DNA coding sequences to generate mRNAs and RNA Pol III transcribes the 5S ribosomal RNA (rRNA) and transfer RNA (tRNA) genes. Higher plants have two extra RNA polymerases, RNA Pol IV and Pol V that function mainly in gene silencing and their catalytic subunits, unlike Pol II, are not required for viability.⁹ The largest and the second largest catalytic subunits of these two polymerases have been identified and characterized. The other components of the multisubunit complexes of the two polymerases have been recently identified and are yet to be molecularly characterized.¹⁰ The largest and the second largest subunits for these two polymerases have been named as NRPD1 and NRPD2 for RNA Pol IV and NRPE1 and NRPD2/NRPE2 for RNA Pol V.⁹

Forward genetics screens and reverse genetics analyses have shown that RNA Pol IV and Pol V are involved in RdDM.¹¹⁻¹⁴ Cytosine methylation and siRNA accumulation have been studied in Pol IV and Pol V mutants. siRNA accumulation is abolished in Pol IV mutants^{11,13} and not in Pol V mutants indicating that Pol IV acting upstream of Pol V and function at different steps in the RdDM. 94% of the loci that produce siRNAs in Arabidopsis are dependent on Pol IV and only one third of these loci require also Pol V function.¹⁵ The components of RdDM machinery including Pol IV and Pol V appeared to localize at the chromosomal loci of the sources of siRNA production and targets of RdDM.¹⁶ The siRNAs produced by RdDM colocalize

with AGO4, DCL3, NRPE1 and RDR2 in cajal bodies.¹⁷ It is worth mentioning that Pol IV is also involved in the production of the natural antisense siRNAs (nat-siRNA) using the dsRNA formed by the overlapping 3' ends of gene pairs.¹⁸ Along with other components of the RdDM, Pol IV was shown to be involved in the spreading of short and long range silencing signals. The substrates and the products of these two polymerases are still not fully known. Possible molecular functions may involve methylated DNA, single stranded RNA or dsRNA as template to produce siRNAs.⁹ However, previous studies suggested that the template of Pol IV is RNA as its function is disrupted by RNase A.¹⁶

Two important recent studies have shown that Pol IV and Pol V evolved from RNA Pol II. The first study led to the identification of the components of Pol IV and Pol V subunits and suggests that these polymerases are Pol II-like polymerases that evolved a specialized role in siRNA production and silencing of transposons, endogenous repeats and transgenes.¹⁰ In the second study, an elegant forward genetic screen has identified a new component referred to as RDM2 (for RNA directed DNA methylation 2).¹⁹ RDM2 shares sequence similarity with the fourth subunit of RNA Pol II (RPB4), but is not an orthologue of RPB4, and evolved to assume different function in Pol IV and Pol V complexes. DRM2 is named as NRPD4/NRPE4 for the Pol IV and Pol V complexes respectively. NRPD4 is required for high level accumulation of siRNA, DNA methylation and transcriptional silencing at transposons and repetitive sequences. This second study confirmed that Pol IV and Pol V exist in a Pol II-like multisubunit complexes and provided a strong evidence that the functionality of Pol IV and Pol V require more subunits in addition to their catalytic subunits.

Current model of RdDM pathway suggests that Pol IV acts first to transcribe methylated DNA to produce the initial transcripts that move into the nucleolus and convert into dsRNA by RDR2. Then, dsRNA is diced by DCL3 and methylated by HEN1 to generate the siRNAs. The siRNAs are loaded onto AGO4 which can bind to NRPE1 and exit the nucleolus together to find and base pair with the Pol V transcripts at the target loci. DRM2 may bind to the siRNA/AGO4/NRPE1 complex in the nucleoplasm as it exits from the nucleolus to methylate the target sequences.

AGO proteins. Argonaute genes are found in bacteria, archaea and eukaryotes and their number varies among different species. Argonaute proteins have been discovered in plants more than a decade ago. Argonautes are the major players in RNA-based gene silencing pathways. Through the involvement of Argonautes in many RNAi-based silencing mechanisms, they contribute to maintain the genome, to produce small noncoding RNAs, to form the heterochromatin and to control RNA stability and protein synthesis. Argonaute proteins are classified into three groups based on their phylogenetic relationships and ability to bind small RNAs. Group 1 members are referred to as AGO proteins and they are capable of binding miRNA and siRNAs. Group 2 members are referred to as PIWI, that bind interacting RNAs (piRNAs) and group 3 members are found only in worms and bind to secondary siRNAs. Arabidopsis and rice genomes contain 10 and

18 group 1 argonaute-like genes respectively.^{20,21} Biochemically, Argonaute proteins are similar to RNase H endonucleases but they use RNA instead of DNA to target the RNA molecules.²² Argonautes contain four distinct functional domains namely the N-terminal, PAZ, MID and PIWI domains.^{23,24} PAZ, MID and PIWI domains have important functions in small RNA pathways. The PAZ domain recognizes the 3' end of small RNAs and the MID domain binds to the 5' phosphate. The PIWI domain exhibits an endonuclease activity that is similar to that of RNase H enzymes.

Plant AGOs are grouped in three clades: in Arabidopsis, clade 1 includes AGO1, AGO5 and AGO10; clade 2 includes AGO2, AGO3 and AGO7; and clade 3 includes AGO4, AGO6, AGO8 and AGO9.²¹ AGO1 was identified more than a decade ago from Arabidopsis mutants that exhibit pleiotropic developmental defects and later it was shown to be involved in post-transcriptional gene silencing.^{25,26} AGO4 and AGO6 were shown to be involved in transcriptional gene silencing. AGO4 was identified first in a forward genetic screen for mutants defective in transcriptional gene silencing of SUPERMAN (SUP) gene locus.²⁷ This study established AGO4 as the AGO that facilitates the transcriptional gene silencing. AGO4 was shown to control locus-specific siRNA accumulation and DNA methylation.²⁸ Moreover, AGO4 was shown to be localized in cajal bodies with other members of the RdDM pathway including 24-nt siRNA, Pol IV, RDR2, DCL3 and NRPE1.¹⁷

The proteins of the RdDM pathway have to work in order (Pol IV, RDR2, DCL3 and AGO4) or otherwise will be mis-localized.¹⁶ In Arabidopsis, AGO4 was shown to interact with NRPE1.⁷ The NRPE1 C-terminal domain forms a complex with AGO4 and siRNA. AGO4, siRNA and NRPE1 complex binds to target sequence through the siRNA:target base pairing. AGO4 may recruit other RdDM components to cause methylation or use its catalytic activity to generate secondary siRNAs to reinforce silencing.²⁹ AGO6 was identified in a forward genetic screen for second site suppressors of the transcriptional gene silencing (TGS) in *ros1-1* mutant.³⁰ Repressor of gene silencing 1 (ROS1) prevents hypermethylation and *ros1-1* mutants exhibit hypermethylation and transcriptional gene silencing of the luciferase reporter gene. *ago61ros1* double mutant exhibits a reactivation of the transcriptionally silent luciferase gene. However, transgene reactivation was stronger in *ago4-1ros1-1* double mutant. AGO6 was shown to be important for the accumulation of specific heterochromatin siRNAs and its functions are partially redundant with that of AGO4.

DNA cytosine methyl transferases. Up to 50% of cytosines are methylated in the Arabidopsis genome. DNA methylation is an important epigenetic modification. In addition to the primary DNA sequence, the chromatin organization plays a major role in determining the gene transcription. The genome is composed of euchromatic and heterochromatic regions. Euchromatin is less compacted and accessible for the transcription machinery and leads to gene expression. However, heterochromatin is highly compacted and renders the DNA not accessible to the transcription machinery and hence leads in general to gene repression. The heterochromatin regions are determined in part by the

methylation of cytosine nucleotides. Hence, DNA methylation leads to gene repression and demethylation leads to gene expression. For example, drastic loss of methylation results in massive reactivation of transposons.³¹⁻³³ Moreover, the control of gene expression by DNA methylation depends on the position of the methylation marks relative to the gene. Methylation in the promoter sequences tends to repress gene expression. However, body methylated genes are expressed in moderate to high levels and may lose their tissue specificity.³⁴ The loss of gene body methylation does not lead to significantly higher gene expression and it might help fine tune the expression in response to developmental or environmental stimuli. However there are exceptions to this rule for example SUPERMAN (SUP) and AGAMOUS gene body methylation causes transcriptional repression because important regulatory regions reside in gene body.^{35,36}

The methylation of cytosine nucleotides in the genome is catalyzed by cytosine methyltransferases. The methylation of cytosine nucleotides determines the extent of heterochromatin and hence the level of gene expression. Cytosine nucleotides are methylated at the 5' position of the pyrimidine ring. This methylation reaction is catalyzed by cytosine methyltransferases by transferring the methyl group from S-adenosyl methionine (SAM) onto the 5' position of the pyrimidine ring generating 5-methyl cytosine (5-me C). The methylation mark on the DNA can serve to attract methyl binding proteins (MBP) that may function as a platform to recruit other chromatin modifying and remodeling complexes. These complexes may function in heterochromatin formation which is repressive for gene expression due to the inaccessibility of the gene regulatory sequences to the transcriptional machinery. Several cytosine methyltransferases are present in the genomes of prokaryotes and eukaryotes.

Cytosines can be found in three sequence contexts; 5' CpG 3', 5' CpHpG 3' and 5' CpHpH 3'. Cytosine in CpG and CpHpG are called symmetric cytosines and in CpHpH is called asymmetric cytosine. After each round of DNA replication during cell division, each daughter cell has a hemimethylated DNA (one methylated parental strand and one newly synthesized and unmethylated strand). For the CpG and CpHpG symmetric methylation sequences, the methylation can be established on the unmethylated strand by maintenance methyltransferases based on the information from the old methylated strand. This methylation can occur in the absence of the original methylation signal. General observations on methylation patterns have been made from methylation profiling studies on the whole genome scale. First, methylation is targeted to transposable elements and repeat rich sequences like centromeric repeats and ribosomal DNA sequences (rDNA). This explains that genomes that are rich in repeat sequences exhibit higher methylation relative to genomes with less repeat sequences (25% of cytosines are methylated in maize compared to 6% in Arabidopsis).^{37,38} Second, methylation at CpG loci is highly abundant followed by CpHpG and CpHpH loci respectively. It should be noted that cytosines occur primarily at CpG dinucleotides in the mammalian genome. However, in plant genomes, cytosine methylation occurs at all sequence contexts indicating a rich repertoire

of regulatory methylation machineries and complex regulatory mechanisms.

A typical cytosine methyltransferase contains four important domains; a binding domain for SAM, a binding site for the DNA target, a catalytic domain that catalyzes the methyl transfer reaction and a genome targeting domain.^{39,40} Three major cytosine methyltransferase classes have been characterized in plants.⁴¹ Methyltransferase 1 (MET1), a plant homologue of the mammalian DNMT1, a chromomethyltransferase (CMT) and domains rearranged methyltransferase (DRM). MET1 is the major methylation maintenance enzyme at the CpG dinucleotides. CpG methylation is strongly reduced in *met1* mutants.³ *met1* mutant exhibits morphological defects such as delayed flowering and reduced size.^{4,42,43} These morphological defects could be traced back to specific loci with reduced methylation like FWA (*flowering wageningen*) gene sequences. Several forward and reverse genetic studies have concluded that MET1 is required to maintain CpG methylation but not non-CpG methylation.⁴³ The residual methylation observed in *met1* mutants is facilitated by other cytosine methyltransferases. The CMT chromomethylase, ubiquitous and specific to plants, is characterized by the presence of a chromodomain motif embedded in the C-terminal domain.⁴⁴ Forward genetic screens have shown that CMTases are involved in non-CpG methylation. A loss-of-function allele of *cmt3* exhibited a strong reduction in CpHpG methylation. *cmt3* mutants do not exhibit severe morphological defects like *met1* indicating that the CpG methylation is the basal and primary pattern of methylation.^{45,46} These observations suggest that non-CpG patterns could provide a secondary level of regulation. Domains rearranged methyltransferase (DRM) is the third class of cytosine methyltransferases in the plant genome. DRM is the plant homologue of the mammalian DNMT3.⁴⁷ DRM contains all of the domains found in the mammalian DNMT3 but in a different arrangement. *drm1* and *drm2* double mutant exhibits no morphological defects and subtle changes in the methylation patterning.^{48,49} This mutant was found to establish a new methylation imprints on FWA and SUP genes respectively. These observations suggest that DRM1 and DRM2 play a role in establishing a new methylation imprint in response to RNA trigger. In conclusion, the RdDM machinery requires different methyltransferases to methylate cytosine in different sequence contexts.^{50,51} MET1 was shown to be dispensable to the initiation of the RdDM but indispensable to maintain the methylation in the absence of RNA trigger. RdDM requires the activity of DRM to establish methylation in all symmetric and asymmetric sequence contexts. *drm2* mutant shows a complete loss of asymmetric methylation and partial loss of CpHpG symmetric methylation the remaining of which is catalyzed by CMT3.⁵²

As RdDM establishes DNA methylation on the non-methylated target sequence, DRM methyltransferase proteins are very important component of RdDM. In Arabidopsis, *drm1drm2* double mutant lacks all types of de novo DNA methylation including transformed tandem repeats and transcribed inverted repeats.^{48,52,53} In the current model of the DRM targeting to specific sequences, transcription of a given gene target produces nascent RNA transcripts that base pair with the 24-nt siRNAs

tethering a complex of chromatin modifying and remodeling enzymes and DRM methyltransferase. This model is supported by the fact that Pol IV/V are important for siRNA directed DNA methylation. In conclusion, DNA methyltransferases play the key role in RdDM and the final methylation patterns and levels are shaped by the activity of both DNA methylases and demethylases.^{54,55} For some genes methyltransferases may be sufficient to generate the exact level and pattern of methylation but some other genes may need the action of demethylases to adjust the level and pattern of methylation. Indeed, DNA demethylases are important in the control of gene activity and to ensure the reversibility and the plasticity of the epigenetic modifications.⁵⁵ It should be noted that the methylation patterns of the genome may not be final throughout the plant life cycle. Methylation patterns can change substantially in response to developmental and environmental signals.⁵⁶

Histone modification and chromatin remodeling proteins.

Chromatin structure has emerged not only to provide a packaging solution of DNA in the nucleus but also as a source of passing a relatively stable genetic information.⁵⁷

As mentioned earlier, highly condensed chromatin, heterochromatin, is repressive for gene expression and open and relaxed chromatin, euchromatin, is permissive for active gene expression.⁵⁸ Most coding sequences are located in euchromatin regions whereas repetitive and transposable elements are located in heterochromatin regions. Usually heterochromatin is subjected to tight and somewhat permanent silencing but euchromatin is subjected to both silencing and activation machineries. Chromatin is composed of nucleosome units which in turn are composed of the DNA helix wrapped around the histone octamers (two copies of each H2A, H2B, H3 and H4). The nucleosomal units exhibit the beads on the string structure and can be connected by histone H1 to generate higher order and compacted chromatin.

Histone proteins can be modified by methylation, acetylation, deacetylation, phosphorylation and ubiquitination to regulate the chromatin structure. Another layer of complexity is added by the fact that lysines in histones can be monomethylated, dimethylated and trimethylated. Also arginines can be monomethylated or dimethylated. It should be noted that histone acetylation is associated with active gene transcription and histone deacetylation is associated with gene repression. However, histone methylation can be associated with either gene expression or repression.⁵⁹ For example, the di or trimethylation of lysine 4 of histone H3 (H3K4me2/3) and the di or trimethylation of lysine 36 of histone H3 (H3K36me2/3) are histone modifications associated with gene expression. However, trimethylation of lysine 27 of histone H3 (H3K27me3) is a histone modification that associates with gene repression. These specific modifications of the histones serve as a code to favor certain chromatin states that contribute to gene expression control.

There is a crosstalk between cytosine methylation and histone modifications on the molecular level. Silent genes were shown to be DNA methylated and deacetylated at histones H3 and H4.⁶⁰ Cytosine methylation can be used as a nucleating signal for histone modifications. The histones of the promoters silenced by RdDM was shown to be deacetylated by AtHDA6.⁶¹ Forward

genetics screens have identified AtHDA6 as an important factor to maintain the silenced state and as an essential component of RdDM that trigger the covalent modifications of histones in a sequence specific manner.⁶² RdDM silenced promoters are reactivated in AtHDA6 loss of function mutants. Silencing of promoters by RdDM involves a cooperation between cytosine methylation and histone deacetylation. Histone methyltransferases including SUVH4/KRYPTONITE (KYP) histone 3 lysine 9 (H3K9) methyltransferase, SUVH5 and SUVH6 are required to maintain the cytosine methylation at CHG and non-CG specific loci respectively.⁶³⁻⁶⁵

Chromatin remodeling factors use the energy from ATP to move, destabilize or slide the nucleosomes. Chromatin remodelers play an important regulatory role in gene transcription and DNA replication, repair and recombination. Some remodelers promote dense nucleosome packaging while other remodelers act to slide or eject nucleosomes to help the access of the transcription machinery. There are three main ATP-dependent chromatin remodeling complexes in all eukaryotes including the switching defective/sucrose nonfermenting (SWI/SNF) ATPases, the initiation switch (ISWI) ATPases and the chromodomain and helicase like domain (CHD) helicases. One of the important components of RNA-directed DNA methylation is the chromatin remodeling factors. A screen for mutants defective in RdDM has identified defective in RNA-directed DNA methylation 1 (DRD1) that belongs to a subfamily of SWI2/SNF2 chromatin remodelers. Arabidopsis genome contains 41 members of the SWI2/SNF2 like chromatin remodeling proteins. DRD1 subfamily of SNF2-like chromatin remodeling proteins is found only in plants and has similarities to Rad54, ATRX and JBP2 subfamilies. DRD1 functions mainly in RNA de novo DNA methylation of the cytosines in all sequence contexts in the target promoter. Interestingly, more CG methylation was retained in the target promoter in the *drd1* mutant than the wild type plants indicating that DRD1 was required for the complete erasure of the CG methylation from the target promoter.⁶⁶ DRD1 plays a dual role, in the presence of RNA trigger it acts to help methylation and in the absence of RNA trigger it helps to erase the methylation.

In addition to DRD1, decrease in DNA methylation (DDM1) is a chromatin remodeling helicase that was shown to maintain the methylation at CpG and non-CpG sequences. *ddm1* mutant exhibits morphological variations that relate to methylation changes at some target loci.⁶⁷ *ddm1* mutant, similar to *met1*, exhibits demethylation of FWA and late flowering phenotype. *ddm1* mutant exhibits reactivation of transposons due to loss of CpG and non-CpG methylations.⁶⁸ As DDM1 exhibits chromatin remodeling activities in vitro and taking into account the previous observations it can be concluded that DDM1 alone or likely in a complex could remodel the chromatin to allow the access of the DNA methylation machinery in the heterochromatin region. Several loci were shown to maintain DNA methylation in the absence of DDM1 including SUP and PAI1-PAI4 and NOSpro inverted repeats.⁶⁹ The loci of these sequences may be available to the methylation machinery and does not require the function of DDM.⁷⁰ Intriguingly, unlike *met1* and *cmt3* mutants, in the *ddm1* mutants the demethylated loci including the centromere

repeats and rDNA repeats remain demethylated for two generations after crossing with the wild type indicating that *ddm1* disrupts a remethylation machinery. It appears that constitutive heterochromatin regions remain repressed and silent and chromatin regions that can change between euchromatin and heterochromatin states can either be active or repressed.⁷¹ This indicates that the methylation/demethylation machineries act on facultative heterochromatin whereas methylation machineries of DNA and chromatin act on constitutive heterochromatin. As RdDM is known to interact with DNA glycosylases and disturbs the DNA methylation it may be acting preferentially on facultative heterochromatin.

Sources and targets of RdDM. The 24-nt-long siRNAs can be generated from dsRNAs from various sources. These sources include RNA Pol II transcription of inverted repeats or overlapping transcription of a single transcript, replication of ssRNA viruses, transposons, aberrant RNAs and ssRNAs generated by Pol V from methylated DNA sequences. Theoretically, all of the sources can serve as targets for the RdDM. Moreover, all the DNA sequences complementary to the isolated siRNAs could be targeted for methylation. Genome wide methylation and transcript profiling have been used to identify the endogenous targets of the RdDM pathway in wild type and mutant backgrounds. Intergenic regions, transposons, repeats and plant genes in euchromatin were identified as major targets of RdDM.^{11,13,14}

In a screen of *drd1* and *Pol V* mutants to identify targets, transposons were found to be upregulated. The reactivation of soloLTR of *Copia*-type LTRCO family was associated with reduced cytosine methylation and increased expression of neighboring sequences. The chromatin state affects the reactivation of targets. Targets that are strongly upregulated were found to reside in euchromatin where targets with weak reactivation reside in heterochromatin. The transposons soloLTR, LTR1 and LTR3 have similar DNA sequence but their reactivation is different according to their chromatin location. soloLTR is highly derepressed because it resides between heterochromatin and euchromatin and carries negligible H3K9me2. However LTR1 and LTR3 reside in heterochromatin and are modified by H3K9me2. It should be noted, however, that not all transposons are regulated by RdDM and alternate mechanisms may exist to control their expression.^{72,73} As the RdDM modification is reversible and the potential targets reside in close proximity to genes in the euchromatin it is likely that RdDM plays a much more significant role, than thought before, in gene expression control in response to developmental or environmental stimuli.

Demethylation Machinery

The final DNA methylation pattern of the genome, in a given tissue or at developmental stage, is determined by the action of DNA methyltransferases and DNA demethylases. Specific de novo DNA methylation that produces a desirable methylation pattern may not require the action of DNA demethylases. However, promiscuous de novo DNA methylation requires the activity of DNA demethylases to produce a desirable pattern.⁵⁵ Passive and active methylation can occur in the genome. DNA

is hemimethylated after DNA replication and the newly synthesized DNA strand requires the action of the maintenance DNA methyltransferases to maintain the methylation pattern. Passive demethylation occurs when the maintenance DNA methyltransferases fail to methylate the newly synthesized strand. Active DNA demethylation involves an enzymatic action of demethylases that remove the methyl group from 5-methyl cytosine (5-mC). In Arabidopsis, several proteins of DNA glycosylases with demethylase activity have been identified which belong to DEMETER family.^{54,74-77} These proteins are repressor of silencing (ROS1), DEMETER (DME), DEMETER like 2 (DML2) and DEMETER like 3 (DML3). The demethylases may use one of several proposed mechanisms to achieve the DNA demethylation. The most well-studied and confirmed mechanism is a base excision repair that is initiated by 5-mC DNA glycosylases. The glycosidic bond between 5-mC and the deoxyribose is cleaved by the DNA glycosylase creating abasic AP site. The deoxyribose at the AP site is removed by the action of AP endonuclease creating a gap that is filled by the DNA polymerase and sealed by DNA ligase. These reactions result in the replacement of the 5-mC by unmethylated cytosine. Other proposed mechanisms of DNA demethylation that awaits more confirmation and replication include nucleotide excision repair and hydrolysis, 5-mC deamination coupled with G/T mismatch repair and oxidation demethylation.

DNA demethylases in mammals have been implicated in early development, memory formation, immune responses and tumorigenesis.⁷⁸⁻⁸⁰ In plants, DNA demethylases function to prevent the RdDM of transgenes and endogenous genes, regulate imprinting and transposons and involve in 5S rDNA chromatin decondensation. Repressor of silencing 1 (ROS1), a DNA glycosylase/lyase, functions to maintain the expression of a transgene and its homologous endogenous gene by preventing the promoter methylation through the RdDM pathway.⁵⁴ Two ROS1 homologues, DML2 and DML3, have been identified in Arabidopsis.^{81,82} A genomic tiling microarray approach was applied to wild type and demethylases triple Arabidopsis mutant (*ros1, dml2, dml3*). 180 loci were shown to be demethylated and more than 80% of which lie in genic regions indicating that demethylation protects the genes from deleterious methylation patterns.⁸³ The expression of several maternally imprinted genes require the action of demethylases. DEMETER (DME) is a DNA demethylase that controls the expression of the maternally imprinted MEDEA (MEA) polycomb group gene in the endosperm tissue. The disruption of the demethylase activity leads to the silencing of the MEA gene in the endosperm and consequently to impaired seed development. Maternally imprinted genes such as FWA and FIS2 (*fertilization independent seed 2*) are controlled by demethylases by similar mechanisms.^{84,85} Several studies have indicated that demethylases play a central role in keeping some of the transposons in a dynamic state and maintaining a basal level of expression.^{54,86} It is quite important to keep basal levels of expression of some transposons especially in large plant genomes because they play central roles in maintaining genome structure and architecture and in inducing genetic variation.^{87,88}

Plants subjected to stress show less methylation profiles irrespective of DNA replication which indicate the involvement of active DNA demethylation. Methylation of the coding region of tobacco glycerophosphodiesterase gene were shown to be greatly reduced when plants are subjected to salt, drought and oxidative stresses.⁸⁹ DNA demethylases function in chromatin decondensation to permit gene expression at specific loci. The 5S rDNA repeats in the centromeric heterochromatin are silenced by RdDM and chromatin remodeling proteins. ROS1 mediates active DNA demethylation and decondenses the 5S rDNA chromatin to permit gene expression during the periods of high protein synthesis demand. The decondensation is reversed by the RdDM machinery. This condensation and decondensation of 5S rDNA allows plants to adjust to growth signals.⁹⁰

In plants, it has been shown that the levels and activities of the demethylases is tightly regulated and coupled to the levels and activities of methyltransferases. In *met1* mutant genome methylation is significantly reduced and the level of ROS1 expression is undetectable.⁹¹ Also, in mutants of the key components of RdDM (*npr1a, rdr2, dcl3, drm2*) the expression of ROS1 is very low.⁹² It is likely that sensors of the methylation activities exist in the genome and relay signals to control the expression of ROS1, DML2 and DML3.⁹³ ROS3, an RNA-binding protein and a component of ROS1 demethylation machinery at some target loci was shown to be regulated also by the methylation levels. ROS3 expression is enhanced in *ros1* mutant and ROS1 expression is enhanced in *ros3* mutant because of increasing methylation at some target loci in *ros1* and *ros3* mutants.⁹⁴ The interplay between methylation and demethylation machineries is used to up or downregulate specific genes or to fine tune the epigenetic states in response to developmental or environmental stimuli. The plasticity of the plant epigenome is maintained by this dynamic control of methylation and demethylation machineries.⁵⁵

Functions of The RdDM

Methylation is very important to plant life. In fact, without the methylation epigenetic modification plant life may not be possible. Methylation is very important in many aspects of plant growth, development, variation, responses to biotic and abiotic stresses and genome stability. As plants rely heavily on the epigenetic changes for their development and responses to the environment and due to their sessile nature, plants may need to assume several epigenomes to control the global expression pattern at any given point in their life cycle.⁹⁵ DNA methylation, histone modifications and chromatin remodeling factors contribute to produce several epigenomes with different expression patterns in response to growth, developmental and environmental stimuli. This notion is supported by the fact that plant genomes contain the largest numbers of chromatin regulatory proteins (e.g., more than 500 in Arabidopsis), DNA and histone methyltransferases. These several epigenomes could produce a long and/or short cellular memories to adjust the plant responses to different internal or external stimuli.

Development. Mitotic epigenetic inheritance through DNA methylation, chromatin modification and remodeling is of

utmost importance to the development of eukaryotes. As cells differentiate to new cell types or tissues and respond differently to different intercellular and environmental signals, the information in the almost static genome sequence may not be enough. Cellular memory comes into play to activate or silence genes. This is of particular importance to plant cells which, unlike mammalian cells, do not terminally differentiate and remain totipotent throughout their life. I will discuss the involvement of several RdDM key players in the development. The morphological phenotypes of mutants of these key players have established the connection of their roles to development. For example, FLOWERING LOCUS C (FLC) is the key gene that mediates the vernalization responses. Vernalization is a cold treatment to seeds that induces epigenetic modification that can be maintained through mitotic cell divisions but not from one generation to the next. Vernalization treatment leads to the repression of FLC by histone methylation.^{96,97} The repression of FLC was shown to be mediated by a 24-nt siRNA.⁹⁸ The 24-nt siRNA is absent from Pol IV, RDR2 and DCL3 mutants. These observations suggest, yet to be rigorously tested, that the repression of FLC might be mediated through RdDM pathway.

Pol IV and Pol V subunits may not be essential for the viability of plants. However, they seem to play important RdDM dependent roles in the development. *nrdp1* and *nrdp2* mutants exhibit a late flowering phenotype when grown under short day conditions as observed in *rdr2*, *dcl3* and *ago4* mutants.^{11,14} The interaction between two alleles that gives rise to a meiotically heritable change in gene expression of one allele is called paramutation. MOP1 (mediator of paramutation 1) and RMR1 (required to maintain repression 1) have been identified in maize by forward genetic screens.^{99,100} MOP1 and RMR1 were found to be orthologues of RDR2 and SNF2 subfamily member DRD1 respectively. *mop1* mutant of maize exhibits delayed flowering, reduced size and spindley and barren stalks.¹⁰¹

Several forward genetic screens and reverse genetic analysis studies have shown that plant methyltransferases play a major role in plant development. *met1cmt3* and *met1drm1drm2* mutants exhibit much stronger developmental phenotype than that of the *met1* mutant alone. These observations indicate a functional redundancy between DNA methyltransferases in development. Differential DNA methylation controls the activities of MEA, FIS2 and FWA imprinted genes in paternal and maternal alleles. The methylation of these genes is mediated by MET1.^{84,85,102} These genes are methylated and silenced as a default state. However, the maternal alleles are hypomethylated by the action of demethylases. The imprinting of these genes is catalyzed by the DME DNA demethylase

DECREASE IN DNA METHYLATION 1 (DDM1), a chromatin remodeling ATPase, is involved in CG and non-CG methylation and heterochromatin silencing.¹⁰³ The *ddm1* mutant exhibits several developmental abnormalities. The *ddm1* induced loss-of-function of BONASI (bns) phenotype was due to the hypermethylation and accumulation of small RNAs. The methylation was shown to be dependent on the presence of a retrotransposon in the 3' region. The presence of this retrotransposon could generate the trigger for the epigenetic modification pathway. The

putative histone demethylase increase in BONSAI methylation 1 prevents hypermethylation by protecting genes from CHG methylation.¹⁰⁴⁻¹⁰⁶ In these studies the IBM was shown to be required for the proper development and may have thousands of gene targets.

Stress responses. Epigenetic stress memory provides the necessary information for the cellular machineries to assume the right epigenome at the right time under stress conditions.¹⁰⁷ The reversibility of the epigenetic changes is crucial for the plant adaptation under different stress conditions. Plants tend to slow their growth to conserve energy as a response to stress so that they do not run the risk of dying.¹⁰⁸ Ribosome biogenesis, a massive energy consuming process, and the rRNA transcription increase in response to increasing growth to support the protein synthesis demand. However, under slowed growth conditions the rate of rRNA transcription and ribosome biogenesis decreases as a cellular economic regulatory mechanism to conserve energy. The rRNA transcription is controlled by a DNA methylation and histone modification epigenetic switch.¹⁰⁹ This rRNA transcription epigenetic switch may be controlled by stress signals.

The reactive oxygen species (ROS) generated by salt stress induces the expression of SRO5 and P5CDH genes which leads to the production of a 24-nt natural small interfering RNAs (nat-siRNA) from their complementary sequences. The formation of 24-nt nat-siRNAs target the P5SDH messages for cleavage and this process is dependent on the NRPD.⁸ Accordingly, the production of nat-siRNAs downregulates the expression of the P5CDH leading to the accumulation of proline which plays an important role in plant stress tolerance.

Differential expression levels of maintenance and de novo DNA methyltransferases was observed in rice plants under cold and salt stress.¹¹⁰ Analysis of the methylation patterns in the pea root tips subjected to water stress has shown that hypermethylation is induced by water deficit and that the hypermethylation at the CCGG sequence occurs in specific DNA targets.¹¹¹ Mild osmotic stress induces reversible hypermethylation of the CpCpG trinucleotides in tobacco cell suspension cultures.¹¹² In this system, there was no changes in the CpG methylation within the CCGG motif. The CG dinucleotides appeared to be fully methylated under normal physiological conditions. The genome of the tobacco cells grown in suspension culture seems to adjust its responses to osmotic stress through de novo methylation of DNA. Aluminum stress induces the transcription of a glycoposphodiesterase-like protein (NtGPDL). Methylation and demethylation have been observed only in the coding region of NtGPDL. In response to stress, the CG dinucleotide is selectively methylated. Salt, cold and oxidative stresses induce the same demethylation patterns indicating that the gene body methylation may affect the chromatin structure and gene expression of NtGPDL gene.⁸⁹

The activity of the transposable elements is known to be regulated by reversible methylation. The regulation of rice retrotransposon can affect the expression patterns of the flanking gene sequences. Changes of the methylation patterns of the retrotransposon are induced by stress and show tissue specificity.¹¹³ The regulation of the transposable elements methylation patterns could provide another layer of controlling gene expression in

response to environmental stress. AtCopeg1 is a member of the *Copia*-like retrotransposons that shows another example of the involvement of retrotransposons in stress responses. AtCopeg1 is expressed in a tissue specific manner and its expression level is modulated by hormones and stress signals.¹¹⁴

Genome stability. Transposons have been discovered because of their propensity to cause damage to the chromosomes. Most of the eukaryotic genomes have a high number of transposons that, if activated, can cause high levels of mutations. However, most of the transposons are quiescent or silenced through epigenetic mechanisms. One of the important functions of RdDM is to target transposons to protect the genome from their deleterious effects. Indeed, siRNAs from transposons have been detected in Arabidopsis and tobacco indicating that transposons are major sources and targets of RdDM. Transposons can also affect the transcription of host genes. For example, FWA gene transcription is inhibited due to the presence of short interspersed nuclear element (SINE) transposon in the promoter of the gene.¹¹⁵ But how the RdDM machinery recognizes and acts upon the transposons not the host genes? Transposable elements may carry some features that make them more distinguishable from other sequences and hence more susceptible to the RdDM. Several lines of evidence have pointed out that transposable elements could possess long untranslated regions or specific secondary structures that make them more recognizable to the methylation machinery. However, several studies in maize and Arabidopsis have indicated that nothing intrinsic to the transposons sequences provides recognition features. Active CACTA transposon in a mutant background can remain active after being introduced in wild type background for several generations.¹¹⁶ The recognition, however, could be due to the fact that transposons produce aberrant RNAs that can be detected by surveillance system. The aberrant RNAs produce dsRNAs which act as precursors to produce the siRNAs and target the RdDM machinery to the transposons sequences in the genome. Indeed, the aberrant RNAs produced from transposons provide a self-reinforcing silencing mechanism.

In Arabidopsis, the CACTA class transposon is methylated at CG and non-CG sequences and almost silent in wild type plants. *met1* and *cmt3* single mutations can lead to partial activation with no mobility of the transposon. However, *met1cmt3* double mutant leads to full activity and mobility of the transposon.¹¹⁷ This indicates that CG and non-CG methylation is required for full silencing of the transposon. It is worth emphasizing, that transposons are methylated at the ends as well as at the middle of their sequence. These methylations features suggest that suppression of transposons activities could be achieved not only through the transcriptional silencing but also by blocking the access of transposases to the ends.¹¹⁷

Transgenerational Inheritance in Plants

Epigenetic silent information is maintained throughout mitosis and meiosis over generations. The heritability of these silencing information led to the concept of epigenetic alleles. Epigenetic alleles or epialleles share the same DNA sequence but exhibit

differential expression due to the epigenetic marks. Several intriguing examples of epialleles have been characterized in Arabidopsis, the epialleles of SUPERMAN is involved in floral development.¹¹⁸ Epialleles can also interact in cis or trans to cause heritable changes. These epiallelic interactions give rise to concepts like paramutation and imprinting. In paramutation, the interaction of two alleles give rise to a meiotically heritable change of one allele. Maize B1 locus, encodes a transcription factor and controls the anthocyanin pigment, provides an excellent example of paramutation. The two alleles of a paramutagenic B' and paramutable B-1 interact. In a heterozygous background, B' epiallele converts B-I epiallele into B' epiallele. Extensive studies have characterized a 7 repeat region 100 kb upstream of the B locus as a control region for this paramutagenic interaction. A genetic suppressor analysis of paramutation has identified a mediator of paramutation (Mop1) as an RNA dependent RNA polymerase 2. This clearly links the paramutation to siRNA machinery. In imprinting, as discussed earlier, the two alleles have differential expression depending on the parent of origin. For example, MEDEA (MEA) is a maternally imprinted gene in the seed endosperm, the paternal allele of MEA is silent while the maternal allele is expressed.

The alleles of epigenetic states and their associated phenotypes were shown to be inherited through meiotic cell division over several generations. This transgenerational heritability was shown to be dependent on the accurate propagation of DNA methylation at the CpG dinucleotides.^{92,119-121} Mutants defective in the maintenance of CG methylation show severe developmental abnormalities.¹²² Loss of function mutation of MET1 suppresses the activity of DNA demethylases and leads to altered RdDM and redistribution of methylation marks. It is likely that the CpG methylation coordinates and stabilizes the epigenetic memory of transgenerational inheritance in Arabidopsis.⁹²

An efficient remethylation mechanism that corrects defects of transgenerational DNA methylation has been shown to operate in the Arabidopsis genome.¹²³ The remethylation has been shown to be facilitated by RNAi. When the RNAi machinery is compromised the selective and progressive remethylation is impaired. Some repeat elements are efficiently targeted by RNAi-dependent de novo methylation. The elucidation of this RNAi-based mechanism in protecting the genome against the loss of transgenerational marks points to key controls of epialleles in variation and adaptation.

Some plants are very amenable to epigenetic inductions where others are resistant. For example, rapid and heritable changes were induced in the inbred flax variety Sormont Cirrus (PI) by growth media and stress conditions.¹²⁴ DNA methylation and histone modifications were suggested as control mechanisms that cause these epigenetic modifications.¹²⁵ Plants subjected to stress and poor growth conditions exhibit progressive epigenetic modification over generations, which leads to stress tolerant phenotype. Moreover, plants that grow in very rich growth medium and optimum conditions also exhibit progressive epigenetic modifications over generations that lead to increased plant size and or yield. However, these plants may lose some parental stress

tolerance epigenetic modifications and become more susceptible once exposed again to adverse environmental conditions.

The extensive genotypic and phenotypic changes that occur in flax in one generation provide a rich resource to study the epigenetic mechanisms and the points of control of such epigenetic modifications. Using this system, important questions like why certain varieties are sensitive to epigenetic inductions and others are resistant? what are the key controls that can unleash the batteries of epigenetic modifications? and how to engineer plants to be sensitive to epigenetic modifications for desirable traits? Such questions might be answered using the information from this system or similar system.

Concluding Remarks

For plants to thrive and adapt in a dynamic and changing environment, Mendelian inheritance may not be sufficient. Epigenetics adaptations to fluctuations in growth conditions and environmental cues help fine tune the epigenome. This epigenome fine tuning helps plants endure for many generations under limited nutrition or stressful conditions. However, plant epigenetic mechanisms appear to be remarkably complex when compared to mammals. DNA methylation is considered to be the most important epigenetic mark. The RdDM serves multiple functions including genome defense, imprinting, paramutation and gene expression control in response to growth, developmental or stress signals. The RdDM involves the production of 24-nt long siRNAs that can be produced from heterochromatin repetitive sequences, inverted repeats or overlapping transcription. The siRNAs biogenesis takes place in the nucleolus and siRNAs are loaded onto AGO4 which interacts with NRPE1. siRNA/AGO4/NRPE1 complex exits the nucleolus to the nucleoplasm to recruit the DRM1/2 de novo methyltransferases to the target DNA sequences. The hallmark of the RdDM is that cytosine methylation occurs in all sequence contexts (CpG, CpHpG and CpHpH). The sources and targets of RdDM appears to reside with high density in heterochromatic regions of transposons and repetitive sequence and with less density in euchromatic regions.

I am grateful to Prof. Jian-Kang Zhu, the director of the center for plant stress genomics & technology at KAUST, for his helpful discussions and valuable comments on the manuscript. I also thank Prof. Ray Bressan for reading the manuscript and insightful discussions.

References

1. Wassenegger M, Heimes S, Riedel L, Sanger HL. RNA-directed de novo methylation of genomic sequences in plants. *Cell* 1994; 76:567-76.
2. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 1992; 69:915-26.
3. Finnegan EJ, Peacock WJ, Dennis ES. Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proc Natl Acad Sci USA* 1996; 93:8449-54.

4. Kankel MW, Ramsey DE, Stokes TL, Flowers SK, Haag JR, Jeddloh JA, et al. Arabidopsis MET1 cytosine methyltransferase mutants. *Genetics* 2003; 163:1109-22.
5. Bender J. DNA methylation and epigenetics. *Annu Rev Plant Biol* 2004; 55:41-68.
6. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002; 16:6-21.
7. El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, et al. Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. *Genes Dev* 2007; 21:2539-44.
8. Chinnusamy V, Zhu JK. RNA-directed DNA methylation and demethylation in plants. *Sci China C Life Sci* 2009; 52:331-43.
9. Pikaard CS, Haag JR, Ream T, Wierzbicki AT. Roles of RNA polymerase IV in gene silencing. *Trends Plant Sci* 2008; 13:390-7.
10. Ream TS, Haag JR, Wierzbicki AT, Nicora CD, Norbeck AD, Zhu JK, et al. Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. *Mol Cell* 2009; 33:192-203.

The heterochromatic targets may be kept silent permanently but the euchromatic targets may exhibit a reversible silencing and hence they are subjected to the actions of methyltransferases and demethylases.

DRM1/2 are the primary de novo methyltransferases that execute the RdDM. However, MET1 and CMT3 methyltransferases were found to be important for de novo methylation at many targets. This could be due to, at least in part, that the efficient symmetric methylation of CpG and CpHpG leads to efficient de novo methylation. It is not clear whether RdDM trigger partial silencing and recruit histone modifying proteins for complete silencing or chromatin modification and remodeling facilitate and lead to RdDM. DNA methylation/demethylation machineries operate in euchromatic regions to insure the reversibility of the repressive marks and the plasticity of the genome.

The core genes of RdDM epigenetic pathway are conserved between monocots and dicots and the mutation in any of these genes exhibits similar phenotypes. However, recent data show that methylation patterns and levels may vary in different genomes. For example unlike the methylation patterns in Arabidopsis, the promoters of the rice genome are enriched in methylation. The epigenetic regulation is affected by the genome architecture and complex and larger genomes may warrant broader roles of epigenetic silencing. There are still a lot to be learned from comparative studies especially from the analysis of dicots vs. monocots and small vs. large genomes. It is expected that RdDM plays a much wider and pronounced roles in complex genomes.

As the technology for whole genome microarrays and high-throughput sequencing and profiling of DNA methylation is available, comparative epigenomic approaches ought to be utilized to identify more RdDM key players and targets across plant genomes and to link the methylation pattern of a given epigenome to growth, developmental and stress cues. This information will help us to learn more about the molecular mechanisms of epigenetics and its contribution to natural phenotypic variation.

Acknowledgements

11. Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA polymerase IV directs silencing of endogenous DNA. *Science* 2005; 308:118-20.
12. Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, et al. Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet* 2005; 37:761-5.
13. Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, Pikaard CS. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* 2005; 120:613-22.
14. Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi MA, et al. Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. *Genes Dev* 2005; 19:2030-40.
15. Mosher RA, Schwach F, Studholme D, Baulcombe DC. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proc Natl Acad Sci USA* 2008; 105:3145-50.

16. Pontes O, Li CF, Nunes PC, Haag J, Ream T, Vitins A, et al. The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell* 2006; 126:79-92.
17. Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, et al. An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* 2006; 126:93-106.
18. Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu JK. Endogenous siRNAs derived from a pair of natural cis-antisense transcripts regulate salt tolerance in Arabidopsis. *Cell* 2005; 123:1279-91.
19. He XJ, Hsu YF, Pontes O, Zhu J, Lu J, Bressan RA, et al. NRPD4, a protein related to the RPB4 subunit of RNA polymerase II, is a component of RNA polymerases IV and V and is required for RNA-directed DNA methylation. *Genes Dev* 2009; 23:318-30.
20. Morel JB, Godon C, Mourrain P, Beclin C, Boutet S, Feuerbach F, et al. Fertile hypomorphic ARGONAUTE (ago1) mutants impaired in post-transcriptional gene silencing and virus resistance. *Plant Cell* 2002; 14:629-39.
21. Nonomura K, Morohoshi A, Nakano M, Eiguchi M, Miyao A, Hirochika H, et al. A germ cell specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. *Plant Cell* 2007; 19:2583-94.
22. Tolia NH, Joshua-Tor L. Slicer and the argonautes. *Nat Chem Biol* 2007; 3:36-43.
23. Hutvagner G, Simard MJ. Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 2008; 9:22-32.
24. Vaucheret H. Plant ARGONAUTES. *Trends Plant Sci* 2008; 13:350-8.
25. Bohmert K, Camus I, Bellini C, Bouchez D, Caboche M, Benning C. AGO1 defines a novel locus of Arabidopsis controlling leaf development. *EMBO J* 1998; 17:170-80.
26. Fagard M, Boutet S, Morel JB, Bellini C, Vaucheret H. AGO1, QDE-2 and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. *Proc Natl Acad Sci USA* 2000; 97:11650-4.
27. Zilberman D, Cao X, Jacobsen SE. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* 2003; 299:716-9.
28. Zilberman D, Cao X, Johansen LK, Xie Z, Carrington JC, Jacobsen SE. Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. *Curr Biol* 2004; 14:1214-20.
29. Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ. Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* 2006; 443:1008-12.
30. Zheng X, Zhu J, Kapoor A, Zhu JK. Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *EMBO J* 2007; 26:1691-701.
31. Lippman Z, Gendrel AV, Black M, Vaughn MW, Dedhia N, McCombie WR, et al. Role of transposable elements in heterochromatin and epigenetic control. *Nature* 2004; 430:471-6.
32. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 2007; 39:61-9.
33. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell* 2006; 126:1189-201.
34. Vaughn MW, Tanurdzic M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, et al. Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biol* 2007; 5:174.
35. Sieburth LE, Meyerowitz EM. Molecular dissection of the AGAMOUS control region shows that cis elements for spatial regulation are located intragenically. *Plant Cell* 1997; 9:355-65.
36. Ito T, Sakai H, Meyerowitz EM. Whorl-specific expression of the SUPERMAN gene of Arabidopsis is mediated by cis elements in the transcribed region. *Curr Biol* 2003; 13:1524-30.
37. Kakutani T, Munakata K, Richards EJ, Hirochika H. Meiotically and mitotically stable inheritance of DNA hypomethylation induced by ddm1 mutation of *Arabidopsis thaliana*. *Genetics* 1999; 151:831-8.
38. Papa CM, Springer NM, Muszynski MG, Meeley R, Kaepler SM. Maize chromomethylase Zea methyltransferase2 is required for CpNpG methylation. *Plant Cell* 2001; 13:1919-28.
39. Bestor TH. The DNA methyltransferases of mammals. *Hum Mol Genet* 2000; 9:2395-402.
40. Cheng X. Structure and function of DNA methyltransferases. *Annu Rev Biophys Biomol Struct* 1995; 24:293-318.
41. Finnegan EJ, Kovac KA. Plant DNA methyltransferases. *Plant Mol Biol* 2000; 43:189-201.
42. Ronemus MJ, Galbiati M, Ticknor C, Chen J, Dellaporta SL. Demethylation-induced developmental pleiotropy in Arabidopsis. *Science* 1996; 273:654-7.
43. Saze H, Mittelsten Scheid O, Paszkowski J. Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nat Genet* 2003; 34:65-9.
44. Henikoff S, Comai L. A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in Arabidopsis. *Genetics* 1998; 149:307-18.
45. Barteel L, Malagnac F, Bender J. Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes Dev* 2001; 15:1753-8.
46. Lindroth AM, Cao X, Jackson JB, Zilberman D, McCallum CM, Henikoff S, et al. Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* 2001; 292:2077-80.
47. Cao X, Springer NM, Muszynski MG, Phillips RL, Kaepler S, Jacobsen SE. Conserved plant genes with similarity to mammalian de novo DNA methyltransferases. *Proc Natl Acad Sci USA* 2000; 97:4979-84.
48. Cao X, Jacobsen SE. Role of the Arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. *Curr Biol* 2002; 12:1138-44.
49. Cao X, Jacobsen SE. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc Natl Acad Sci USA* 2002; 99:16491-8.
50. Jones L, Ratcliff F, Baulcombe DC. RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. *Curr Biol* 2001; 11:747-57.
51. Zaratigui M, Irvine DV, Martienssen RA. Noncoding RNAs and gene silencing. *Cell* 2007; 128:763-76.
52. Cao X, Aufsatz W, Zilberman D, Mette MF, Huang MS, Matzke M, Jacobsen SE. Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr Biol* 2003; 13:2212-7.
53. Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE. RNA silencing genes control de novo DNA methylation. *Science* 2004; 303:1336.
54. Zhu J, Kapoor A, Sridhar VV, Agius F, Zhu JK. The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation patterns in Arabidopsis. *Curr Biol* 2007; 17:54-9.
55. Zhu JK. Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet* 2009.
56. Burn JE, Bagnall DJ, Metzger JD, Dennis ES, Peacock WJ. DNA methylation, vernalization, and the initiation of flowering. *Proc Natl Acad Sci USA* 1993; 90:287-91.
57. Jenuwein T, Allis CD. Translating the histone code. *Science* 2001; 293:1074-80.
58. Richards EJ, Elgin SC. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 2002; 108:489-500.
59. Kouzarides T. Chromatin modifications and their function. *Cell* 2007; 128:693-705.
60. Dobosy JR, Selker EU. Emerging connections between DNA methylation and histone acetylation. *Cell Mol Life Sci* 2001; 58:721-7.
61. Aufsatz W, Mette MF, van der Winden J, Matzke M, Matzke AJ. HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. *EMBO J* 2002; 21:6832-41.
62. Matzke MA, Birchler JA. RNAi-mediated pathways in the nucleus. *Nat Rev Genet* 2005; 6:24-35.
63. Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 2002; 416:556-60.
64. Malagnac F, Barteel L, Bender J. An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J* 2002; 21:6842-52.
65. Ebbs ML, Barteel L, Bender J. H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. *Mol Cell Biol* 2005; 25:10507-15.
66. Huettel B, Kanno T, Daxinger L, Bucher E, van der Winden J, Matzke AJ, Matzke M. RNA-directed DNA methylation mediated by DRD1 and Pol IVb: a versatile pathway for transcriptional gene silencing in plants. *Biochim Biophys Acta* 2007; 1769:358-74.
67. Vongs A, Kakutani T, Martienssen RA, Richards EJ. *Arabidopsis thaliana* DNA methylation mutants. *Science* 1993; 260:1926-8.
68. Jacobsen SE, Sakai H, Finnegan EJ, Cao X, Meyerowitz EM. Ectopic hypermethylation of flower-specific genes in Arabidopsis. *Curr Biol* 2000; 10:179-86.
69. Barteel L, Bender J. Two Arabidopsis methylation-deficiency mutations confer only partial effects on a methylated endogenous gene family. *Nucl Acids Res* 2001; 29:2127-34.
70. Aufsatz W, Mette MF, van der Winden J, Matzke AJ, Matzke M. RNA-directed DNA methylation in Arabidopsis. *Proc Natl Acad Sci USA* 2002; 99:16499-506.
71. Schoft VK, Chumak N, Mosiolek M, Slusarz L, Komnenovic V, Brownfield L, et al. Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. *EMBO Rep* 2009; 10:1015-21.
72. Rangwala SH, Richards EJ. Differential epigenetic regulation within an Arabidopsis retroposon family. *Genetics* 2007; 176:151-60.
73. Habu Y, Mathieu O, Tariq M, Probst AV, Smathajitt C, Zhu T, Paszkowski J. Epigenetic regulation of transcription in intermediate heterochromatin. *EMBO Rep* 2006; 7:1279-84.
74. Agius F, Kapoor A, Zhu JK. Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc Natl Acad Sci USA* 2006; 103:11796-801.
75. Gong Z, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, David L, Zhu JK. ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell* 2002; 111:803-14.
76. Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marin MI, Martinez-Macias MI, Ariza RR, Roldan-Arjona T. DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci USA* 2006; 103:6853-8.
77. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, et al. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. *Cell* 2002; 110:33-42.

78. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. *Nature* 2000; 403:501-2.
79. Miller CA, Sweatt JD. Covalent modification of DNA regulates memory formation. *Neuron* 2007; 53:857-69.
80. Reiner SL. Epigenetic control in the immune response. *Hum Mol Genet* 2005; 14:41-6.
81. Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, et al. DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell* 2006; 124:495-506.
82. Arnaud P, Feil R. MEDEA takes control of its own imprinting. *Cell* 2006; 124:468-70.
83. Penterman J, Uzawa R, Fischer RL. Genetic interactions between DNA demethylation and methylation in *Arabidopsis*. *Plant Physiol* 2007; 145:1549-57.
84. Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, et al. One-way control of FWA imprinting in *Arabidopsis* endosperm by DNA methylation. *Science* 2004; 303:521-3.
85. Jullien PE, Kinoshita T, Ohad N, Berger F. Maintenance of DNA methylation during the *Arabidopsis* life cycle is essential for parental imprinting. *Plant Cell* 2006; 18:1360-72.
86. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* 2008; 133:523-36.
87. Bennetzen JL, Ma J, Devos KM. Mechanisms of recent genome size variation in flowering plants. *Ann Bot* 2005; 95:127-32.
88. Feschotte C, Jiang N, Wessler SR. Plant transposable elements: where genetics meets genomics. *Nat Rev Genet* 2002; 3:329-41.
89. Choi CS, Sano H. Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. *Mol Genet Genomics* 2007; 277:589-600.
90. Douet J, Blanchard B, Cuvillier C, Tourmente S. Interplay of RNA Pol IV and ROS1 during post-embryonic 5S rDNA chromatin remodeling. *Plant Cell Physiol* 2008; 49:1783-91.
91. Huettel B, Kanno T, Daxinger L, Aufsatz W, Matzke AJ, Matzke M. Endogenous targets of RNA-directed DNA methylation and Pol IV in *Arabidopsis*. *EMBO J* 2006; 25:2828-36.
92. Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J. Transgenerational stability of the *Arabidopsis* epigenome is coordinated by CG methylation. *Cell* 2007; 130:851-62.
93. Zemach A, Grafi G. Methyl-CpG-binding domain proteins in plants: interpreters of DNA methylation. *Trends Plant Sci* 2007; 12:80-5.
94. Zheng X, Pontes O, Zhu J, Miki D, Zhang F, Li WX, et al. ROS3 is an RNA-binding protein required for DNA demethylation in *Arabidopsis*. *Nature* 2008; 455:1259-62.
95. Zhu JK. Epigenome sequencing comes of age. *Cell* 2008; 133:395-7.
96. Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C. Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 2004; 427:164-7.
97. Sung S, Amasino RM. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature* 2004; 427:159-64.
98. Swiezewski S, Crevillen P, Liu F, Ecker JR, Jerzmanowski A, Dean C. Small RNA-mediated chromatin silencing directed to the 3' region of the *Arabidopsis* gene encoding the developmental regulator, FLC. *Proc Natl Acad Sci USA* 2007; 104:3633-8.
99. Alleman M, Sidorenko L, McGinnis K, Seshadri V, Dorweiler JE, White J, et al. An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* 2006; 442:295-8.
100. Hale CJ, Stonaker JL, Gross SM, Hollick JB. A novel Snf2 protein maintains trans-generational regulatory states established by paramutation in maize. *PLoS Biol* 2007; 5:275.
101. Dorweiler JE, Carey CC, Kubo KM, Hollick JB, Kermicle JL, Chandler VL. mediator of paramutation1 is required for establishment and maintenance of paramutation at multiple maize loci. *Plant Cell* 2000; 12:2101-18.
102. Xiao W, Gehring M, Choi Y, Margossian L, Pu H, Harada JJ, et al. Imprinting of the MEA Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. *Dev Cell* 2003; 5:891-901.
103. Jeddeloh JA, Bender J, Richards EJ. The DNA methylation locus DDM1 is required for maintenance of gene silencing in *Arabidopsis*. *Genes Dev* 1998; 12:1714-25.
104. Saze H, Sasaki T, Kakutani T. Negative regulation of DNA methylation in plants. *Epigenetics* 2008; 3:122-4.
105. Miura A, Nakamura M, Inagaki S, Kobayashi A, Saze H, Kakutani T. An *Arabidopsis* jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. *EMBO J* 2009; 28:1078-86.
106. Saze H, Shiraiishi A, Miura A, Kakutani T. Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. *Science* 2008; 319:462-5.
107. Chinnusamy V, Zhu JK. Epigenetic regulation of stress responses in plants. *Curr Opin Plant Biol* 2009; 12:133-9.
108. Mahfouz MM, Kim S, Delauney AJ, Verma DP. *Arabidopsis* TARGET OF RAPAMYCIN interacts with RAPTOR, which regulates the activity of S6 kinase in response to osmotic stress signals. *Plant Cell* 2006; 18:477-90.
109. Woo HR, Richards EJ. Natural variation in DNA methylation in ribosomal RNA genes of *Arabidopsis thaliana*. *BMC Plant Biol* 2008; 8:92.
110. Sharma R, Mohan Singh RK, Malik G, Deveshwar P, Tyagi AK, Kapoor S, Kapoor M. Rice cytosine DNA methyltransferases—gene expression profiling during reproductive development and abiotic stress. *FEBS J* 2009.
111. Labra M, Ghiani A, Citterio S, Sgorbati S, Sala F, Vannini C, et al. Analysis of cytosine methylation pattern in response to water deficit in pea root tips. *Plant Biol* 2002; 4:694-9.
112. Kovarik A, Matyasek R, Leitch A, Gazdova B, Fulnecek J, Bezdek M. Variability in CpNpG methylation in higher plant genomes. *Gene* 1997; 204:25-33.
113. Kashkush K, Khasdan V. Large-scale survey of cytosine methylation of retrotransposons and the impact of readout transcription from long terminal repeats on expression of adjacent rice genes. *Genetics* 2007; 177:1975-85.
114. Duan K, Ding X, Zhang Q, Zhu H, Pan A, Huang J. AtCopeg1, the unique gene originated from AtCopia95 retrotransposon family, is sensitive to external hormones and abiotic stresses. *Plant Cell Rep* 2008; 27:1065-73.
115. Kinoshita Y, Saze H, Kinoshita T, Miura A, Soppe WJ, Koornneef M, Kakutani T. Control of FWA gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. *Plant J* 2007; 49:38-45.
116. Kato M, Takashima K, Kakutani T. Epigenetic control of CACTA transposon mobility in *Arabidopsis thaliana*. *Genetics* 2004; 168:961-9.
117. Kato M, Miura A, Bender J, Jacobsen SE, Kakutani T. Role of CG and non-CG methylation in immobilization of transposons in *Arabidopsis*. *Curr Biol* 2003; 13:421-6.
118. Jacobsen SE, Meyerowitz EM. Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. *Science* 1997; 277:1100-3.
119. Soppe WJ, Jasencakova Z, Houben A, Kakutani T, Meister A, Huang MS, et al. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. *EMBO J* 2002; 21:6549-59.
120. Tariq M, Saze H, Probst AV, Lichota J, Habu Y, Paszkowski J. Erasure of CpG methylation in *Arabidopsis* alters patterns of histone H3 methylation in heterochromatin. *Proc Natl Acad Sci USA* 2003; 100:8823-7.
121. Mathieu O, Probst AV, Paszkowski J. Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis*. *EMBO J* 2005; 24:2783-91.
122. Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. *Proc Natl Acad Sci USA* 1996; 93:12406-11.
123. Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, Ciaudo C, et al. A role for RNAi in the selective correction of DNA methylation defects. *Science* 2009; 323:1600-4.
124. Cullis CA. Mechanisms and control of rapid genomic changes in flax. *Ann Bot* 2005; 95:201-6.
125. Richards EJ. DNA methylation and plant development. *Trends Genet* 1997; 13:319-23.