DNA methylation and imprinting in plants: machinery and mechanisms

DNA Methylation and Imprinting in Plants: Machinery and Mechanisms

P.R. V. Satyaki[1] and Mary Gehring[1,2]

[1]Whitehead Institute for Biomedical Research, [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA.

Corresponding Author: mgehring@wi.mit.edu

ABSTRACT

Imprinting is an epigenetic phenomenon in which genes are expressed selectively from either the maternal or paternal alleles. In plants, imprinted gene expression is found in a tissue called the endosperm. Imprinting is often set by a unique epigenomic configuration in which the maternal chromosomes are less DNA methylated than their paternal counterparts. In this review, we synthesize studies that paint a detailed molecular portrait of the distinctive endosperm methylome. We will also discuss the molecular machinery that shapes and modifies this methylome, and the role of DNA methylation in imprinting.

Keywords: Gene imprinting, DNA methylation, plants, seeds, endosperm

INTRODUCTION

Imprinting is a phenomenon in which the maternal and paternal alleles of a gene are differentially expressed. Epigenetic modifications to DNA or chromatin, in some cases, direct the preferential activity of one allele over another. Imprinting is widespread among flowering plants (angiosperms), marsupials, and eutherian mammals. The common thread uniting these distant genera is that embryonic development occurs while the embryo is attached to and supported by the mother. A common site of imprinting in animals is the placenta. In plants, imprinting is detected mostly in the endosperm, an analog of the placenta. Endosperm surrounds the embryo and supplies nutrients to it from the maternal parent (**Figure 1**). Imprinting in these tissues might be related to evolutionary pressures that select for imprinted gene expression (Haig 2013, Spencer and Clark 2014).

The endosperm is one product of the double fertilization event that characterizes flowering plant reproduction (**Figure 1**). Briefly, each pollen grain contains two genetically identical haploid sperm cells. Upon reaching an ovule, one sperm fertilizes the haploid egg cell and another sperm fertilizes the diploid central cell, which is adjacent to the sister egg cell. The fertilized egg cell divides to form the diploid embryo. The fertilized central cell gives rise to the triploid endosperm. Although genetically identical except for an additional maternal genome in the endosperm, the two products of fertilization have very different epigenetic states and developmental fates (**Figure 1**). The embryo is the next generation of plant, whereas the endosperm supports the embryo and does not pass on any genetic material. Despite its transient nature, the endosperm is one the most important tissues in the biosphere, providing a significant portion of the world's nutrition (Li and Berger 2012). Approximately 100-300 imprinted genes have been identified in the endosperm of various species (Gehring *et al.* 2011, Hsieh *et al.* 2011, Luo *et al.* 2011, Wolff *et al.* 2011, Zhang *et al.* 2011, Waters *et al.* 2013, Xu *et al.* 2014, Hatorangan *et al.* 2016, Klosinska *et al.* 2016). Plants with disrupted imprinting or disrupted imprinted genes sometimes display defects in endosperm development (Kiyosue *et al.* 1999, Costa *et al.* 2012, Figueiredo *et al.* 2015).

A key molecular controller of imprinting is DNA methylation, or 5-methylcytosine. DNA methylation is often correlated with transcriptional silencing. In the plant kingdom, DNA methylation is found at cytosines in all sequence contexts – CG, CHG and CHH (where H is C,

A, or T). Sequence context is an indication of the origin of the methylation and thus these sites are often considered separately. Methylation is regulated and maintained by a bevy of reinforcing independent pathways that can be broadly divided into two types – the maintenance methylation pathways and the *de novo* methylation pathways (Law and Jacobsen 2010). CG and CHG residues, because of their symmetric nature, are substrates for maintenance methyltransferases, which recognize hemi-methylated DNA after replication and methylate the newly synthesized, unmethylated strand. *De novo* methyltransferases establish DNA methylation in all sequence contexts and are required for maintenance of asymmetric (CHH) methylation. *De novo* methylation is targeted either by small RNAs or by proteins that interact with modified histones. DNA methylation processes are antagonized by active DNA demethylation, which in plants is mediated by a family of 5-methylcytosine DNA glycosylase enzymes (Law and Jacobsen 2010). 5-methylcytosine is removed from DNA by base excision repair, a process essential for establishing imprinted gene expression and for normal endosperm development.

Here we focus on the relationship between differential DNA methylation and imprinted gene expression in plants. We summarize the protein machinery and processes that determine endosperm methylation patterns and outline and assess current models for how DNA methylation regulates the expression of imprinted loci.

THE ALTERED ENDOSPERM METHYLATION LANDSCAPE

In vegetative tissues, like seedlings and leaves, DNA methylation is concentrated in transposable elements (TEs) and repeated sequences, but is also found within the transcribed region of genes. These sequences have very different methylation characteristics. In general, genes contain only CG methylation, which is preferentially located in the 3' portion, away from the transcriptional start site. By contrast, CG and non-CG methylation accumulates throughout the entire portion of transposable elements and repeats. Mutants that disrupt the global patterns of DNA methylation have severe phenotypes (Ronemus *et al.* 1996, Cao and Jacobsen 2002, Zhang and Jacobsen 2006), some of which are associated with the activation of transposable elements, although the precise molecular basis of most phenotypes is not known. The endosperm has a distinct methylation profile compared to other plant tissues (Gehring *et al.* 2009, Hsieh *et al.* 2009, Zemach *et al.* 2010), and is characterized by DNA hypomethylation, as described in detail below.

Recent genome-wide methylation studies have generated DNA methylation data from embryo and endosperm (Gehring *et al.* 2009, Hsieh *et al.* 2009, Zemach *et al.* 2010, Ibarra *et al.* 2012, Pignatta *et al.* 2014, Zhang *et al.* 2014, Wang *et al.* 2015, Klosinska *et al.* 2016). In some cases these data are allele-specific, meaning that the embryo and endosperm profiled is an F_1 derived from two polymorphic parents. Thus two different types of comparisons have been used to gain insights into endosperm methylation. The first comparison is between endosperm and embryo produced from non-polymorphic parents (Gehring *et al.* 2009). In this comparison, as both endosperm and embryo are generated from fertilization by sperm that are assumed to be epigenetically identical, any differentially methylated regions in the endosperm relative to the embryo represent a methylation change on the maternal alleles (**Figure 1**). When data of sufficient depth from inter-strain crosses is available, methylation can be directly compared between maternally and paternally inherited chromosomes (Ibarra *et al.* 2012). This is because sequence polymorphisms between maternal and paternal parents allow the parent-of-origin of methylation to be assigned for a portion of the genome. Both these comparisons produce similar results and have been used to examine the methylomes of endosperm and embryo in *A. thaliana*, *A. lyrata,* rice, and maize (Gehring *et al.* 2009, Hsieh *et al.* 2009, Zemach *et al.* 2010, Ibarra *et al.* 2012, Zhang *et al.* 2014, Klosinska *et al.* 2016). Examining plant species separated by more than 150 million years of evolution (e.g. monocots *vs*. dicots) has ushered in some generalizable rules about endosperm methylation. First, endosperm DNA is generally less methylated than embryo DNA on a genome-wide scale. Critically, within the endosperm, the maternal genome is less methylated than the paternal genome. Maternal genome hypomethylation is restricted to discrete locations in the genome, referred to as differentially methylated regions (DMRs). Each DMR is usually a few hundred base pairs in size. Finally, maternal genome hypomethylation is enriched at or near genes.

Observations in *Arabidopsis* and rice endosperm at multiple developmental stages, and in the gametes before fertilization, indicate that DNA methylation is developmentally dynamic during reproductive development (Xing *et al.* 2015, Moreno-Romero *et al.* 2016, Park *et al.* 2016). In *Arabidopsis* and rice*,* CG hypomethylation is detectable at genic sites in the central cell, like in the endosperm. Yet CG, CHG and CHH methylation at TEs remains fairly high in the central compared to endosperm (Park *et al.* 2016). These data suggest that additional DNA demethylation occurs in the endosperm after fertilization. In rice, a demethylation event 2-3 days

after pollination (DAP) leads to lower genic CG methylation in the endosperm and lower CHH and CHG methylation over TEs (Xing *et al.* 2015). In *Arabidopsis*, the earliest examined endosperm is from 4 DAP (Moreno-Romero *et al.* 2016). At this developmental time point, CG methylation on maternal alleles in the endosperm is already reduced relative to the central cell and CHH methylation is completely absent (Moreno-Romero *et al.* 2016, Park *et al.* 2016). By 6 DAP, CHH methylation returns genome-wide, except on regions of maternal chromosomes that already lacked CG methylation (Ibarra *et al.* 2012, Pignatta *et al.* 2014). This suggests that *de novo* methylation activity in the endosperm after fertilization is more effective on paternally inherited DNA than maternally inherited DNA.

Several species-specific methylation features have also been noted. In *A. thaliana* and *A. lyrata*, CG hypomethylated DMRs are also enriched over TEs, many of which are proximal to genes (Gehring *et al.* 2009, Klosinska *et al.* 2016). Yet even the endosperm methylomes of *A. lyrata* and *A. thaliana* differ in some specific features. In *A. lyrata* endosperm, in contrast to *A. thaliana*, a subset of gene bodies gain CHG methylation on the maternally inherited alleles (Klosinska *et al.* 2016). Other species offer additional variations – castor bean endosperm is also hypomethylated in the CG context, but unlike most other species CHH methylation is not reduced compared to the embryo (Xu *et al.* 2014, 2016). These species-specific methylation patterns are exciting as they relate to possible mechanisms of imprinted gene expression, and suggest that the imprinting machinery or mechanisms could be evolutionarily fluid.

THE MACHINERY UNDERLYING THE METHYLOME

DNA demethylation

The foundation for the DNA methylation landscape in the *Arabidopsis* endosperm is established before fertilization in the female and male gametes, the central cell and the sperm cell. In the *A. thaliana* central cell and later in the endosperm, this landscape is shaped by the antagonistic relationship between the activities of the 5-methylcytosine DNA glycosylase (aka a DNA demethylase) DEMETER (DME) and a host of DNA methyltransferases. DNA glycosylases remove bases from DNA in a process known as base excision repair (BER). The first step in base excision repair is cleavage of the glycosidic bond between the sugar phosphate backbone and the targeted DNA base. This leads to release of the base (in this case, 5mC) and formation of an

abasic (AP) site. Bifunctional DNA glycosylases also have lyase activity and nick the DNA backbone. Biochemical characterization of DME indicates it possesses both glycosylase and lyase activities (Gehring *et al.* 2006). After creation of a nick, AP endonuclease acts on the DNA to leave a 3'-OH group, allowing DNA polymerase to insert a base (C in this case); the sugar-phosphate backbone is then sealed by DNA ligase. DME excises 5-methylcytosine in both CG and non-CG contexts *in vitro* but cleaves hemi-methylated positions more effectively (Gehring *et al.* 2006).

A wild type maternal *DME* allele is required for the development of viable seeds and to establish imprinted gene expression at some loci (Choi *et al.* 2002, Gehring *et al.* 2006). In the absence of *DME*, fertilization occurs and seeds begin to develop, but arrest and die after a few days (Choi *et al.* 2002). Unlike typical DNA glycosylases, DME is a large protein with extensive sequence outside of the DNA glycosylase domain. Three domains are necessary for its 5-methylcytosine DNA glycosylase activity (Choi *et al.* 2002, Mok *et al.* 2010): the DNA glycosylase domain, which has similarity to the HhH-GPD domain found in other DNA BER proteins, and two accessory domains of unknown function (Mok *et al.* 2010). An iron-sulfur cluster (Fe_4-S_4) within the glycosylase domain is necessary for 5-mC DNA glycosylase activity (Mok *et al.* 2010). DME's N terminal region has similarity to histone H1 (Choi *et al.* 2002), but the functional significance of this, if any, is unknown.

In the ovule, expression of *DME* reporter transgenes is detectable in the synergid cells (accessory cells in the ovule that attract the pollen tube for sperm delivery) and the polar nuclei (Choi *et al.* 2002). One of the two synergid cells survives the fertilization process and eventually fuses with the endosperm (Maruyama *et al.* 2015). The polar nuclei fuse to give rise to the mature diploid central cell nucleus, where *DME* expression is high. After fertilization, levels of the *DME* transcript in seeds diminish as the endosperm develops (Choi *et al.* 2002). Consistent with this expression pattern, comparison of *A. thaliana* wild-type and *dme* central cell methylomes shows evidence for discrete DME-dependent demethylation (Park *et al.* 2016).Thus, it is thought that DME's effect on endosperm methylation is primarily caused by its activity before fertilization (**Figure 1**).

DNA methylation has also been profiled in wild type and *dme* mutant endosperm (Gehring *et al.* 2009, Hsieh *et al.* 2009, Ibarra *et al.* 2012). At the genomic level, loss of *DME* generates complex endosperm methylation phenotypes. In *dme* mutants, DNA methylation is retained in the gene flanking regions and transposable elements that are normally CG hypomethylated on the maternal allele, indicating that *DME* is responsible for maternal allele CG hypomethylation in wild type endosperm, consistent with its biochemical activity and expression pattern. However, *dme* mutant endosperm also exhibits CHG and CHH hypomethylation in TEs, a puzzling observation (Gehring *et al.* 2009, Hsieh *et al.* 2009). DNA hypomethylation is the opposite phenotype to that expected when a demethylation pathway is disrupted. However, recent data suggest that this effect is likely a secondary one because *dme* mutant central cells do not exhibit TE hypomethylation (Park *et al.* 2016). CHG and CHH methylation in heterochromatic TEs is maintained by the CMT2 chromomethyltransferase enzyme and by the RNA-directed DNA methylation (RdDM) pathway. RdDM requires accessible chromatin to act (Schoft *et al.* 2009, Gent *et al.* 2014). It could be that DME's demethylation activity makes chromatin more accessible such that active DNA demethylation actually promotes non-CG methylation. This would be consistent with the loss of prominent chromocenters in wild type central cells (Pillot *et al.* 2010, Yelagandula *et al.* 2014). The RdDM machinery may lose access to chromatin in *dme* mutants, giving rise to non-CG hypomethylation.

The patterns of maternal allele demethylation that have been observed in maize and rice are consistent with the activity of a DME-like enzyme, and *DME* orthologs have been detected in monocots (Kapazoglou *et al.* 2013). The putative ortholog of *DME* in rice is *Os01g11900,* also referred to as *ROS1A* (Ono *et al.* 2012). *ROS1A* mutant rice plants exhibit seed phenotypes that are reminiscent of the *Arabidopsis dme* phenotype (Ono *et al.* 2012). In maize, a *DME*-like gene with high endosperm expression has been observed, but a potential role in endosperm demethylation has not been described (Wang *et al.* 2015).

Several other putative factors necessary for DME mediated demethylation have been identified. NAR1/GOLLUM and AtDRE2 are involved in the assembly and maturation of Fe_4-S_4 proteins and mutations in these genes phenocopy *dme* to varying extents (Nakamura *et al.* 2013, Buzas *et al.* 2014). *AtDRE2* mutant genomes are hypermethylated at some DME targets. Both *atdre2* and *nar1* mutants also have defects in the expression of imprinted genes that are regulated by DME.

These results support the idea that the Fe-S protein assembly pathway is important for the formation of a functional DME (Nakamura *et al.* 2013, Buzas *et al.* 2014), consistent with *in vitro* data demonstrating that the Fe-S cluster is necessary for 5-mC DNA glycosylase activity (Mok *et al.* 2010). However, some of the *nar1* and *atdre2* phenotypes could also relate to their role in the maturation of the other Arabidopsis 5mC DNA glycosylases, ROS1, DML2 and DML3, mutations in which affect the chromatin structure of the central cell (Pillot *et al.* 2010).

Plants with mutations in genes that encode proteins that complete DNA repair downstream of DME might also have DNA methylation and seed phenotypes, although the interpretation of any observed phenotypes could be complicated by function in other DNA repair contexts. Double mutants between genes encoding an AP endonuclease and a phosphatase, ZDP, exhibit a seed abortion phenotype like *dme* (Li *et al.* 2015). The polymerase that would fill in AP sites produced by DME activity is unknown. Mutations in *LIG1*, the primary DNA ligase in Arabidopsis, lead to the formation of an under-proliferating endosperm with enlarged nuclei. Interestingly, the *dme* mutation suppresses the *lig1* mutant phenotype; this observation is the basis for the hypothesis that LIG1 completes DNA repair at the sites of DME action (Andreuzza *et al.* 2010).

How DME is targeted to specific genomic sites is an important aspect of DME function that is not well understood. One candidate targeting protein is SSRP1, a histone chaperone (Ikeda *et al.* 2011). *SSRP1* is expressed in the central cell and is necessary for DNA demethylation at a few loci. *ssrp1* mutants, like *dme,* have an endosperm over proliferation phenotype and imprinting defects (Ikeda *et al.* 2011). Another potential candidate that may help target DME is the H1 histone *H1.2* (Rea *et al.* 2012). H1.2 was identified as protein that interacted with DME in a yeast-two hybrid screen. Knockdown of *H1.2* resulted in a loss of imprinting and allele specific expression at two maternally expressed imprinted genes (Rea *et al.* 2012). However, either because of an incomplete knockdown of *H1.2*, or because of a limited role for H1.2 in DME activity, these *H1.2* knockdown seeds show only limited effects on seed development, unlike *dme* mutants (Rea *et al.* 2012). The full extent of SSRP1's or H1.2's contribution to DME function remains unknown.

CG methylation

DNA demethylation mediated by DME counteracts some of the activities of DNA methyltransferases. Mutations in the predominant DNA methyltransferase in Arabidopsis, the CG maintenance methyltransferase MET1, suppress the *dme* seed abortion phenotype when also inherited through the female (Xiao *et al.* 2003). Biochemically, MET1 acts on hemi-methylated DNA produced by DNA replication to convert it to fully methylated DNA. MET1 acts on CG sites but is also required for maintenance of non-CG methylation (Stroud *et al.* 2013) and is expressed in most plant tissues. MET1 activity is also needed maternally, paternally, and in the endosperm for imprinting (Jullien *et al.* 2006, Hsieh *et al.* 2011). However, as assayed by transcriptional reporter genes, *MET1* expression appears low or non-existent in the central cell and is repressed by the MSI1-RB complex (Johnston *et al.* 2008, Jullien *et al.* 2008, 2012) and after fertilization is expressed at low levels in the endosperm (**Figure 2**) (Schmidt *et al.* 2013). There may be synergy between the downregulation of *MET1* and the upregulation of *DME* in the central cell to promote a hypomethylated state. DME cleaves hemi-methylated DNA better than fully methylated DNA (Gehring *et al.* 2006). Reduced *MET1* expression could plausibly lead to a loss of DNA methylation on one strand which, when followed by *DME* activity, would result in complete demethylation. Although this is an intriguing model, recent analysis of the central cell methylome indicates that there is no global loss of CG methylation, as would be expected if maintenance methylation activity was completely compromised (Park *et al*. 2016). Maintenance methylation could also be accomplished by MET1 paralogs (**Figure 2**). Reporter gene studies have shown that *MET2a* and *MET2b* are expressed in the central cell (Jullien *et al.* 2012). *MET2a* is required for the maintenance of methylation at active transposons (Stroud *et al.* 2013, Quadrana *et al.* 2016). A third paralog, *MET3* (*MEE57*), is a paternally expressed imprinted gene (PEG) whose expression is enriched in a sub-region of the endosperm, the chalaza (**Figure 2**) (Jullien *et al.* 2012, Belmonte *et al.* 2013). The roles played by these methyltransferases in the endosperm are unknown.

MET1 orthologs are widely distributed among angiosperms. Loss of MET1 activity in rice is associated with seed lethality (Hu *et al.* 2014, Yamauchi *et al.* 2014). MET1 is recruited to hemi-methylated DNA by the VIM family proteins and loss of VIM protein function phenocopies loss of *MET1* (Achour *et al.* 2008, Nishiyama *et al.* 2013, Shook and Richards 2014). The *A. thaliana* genome encodes six proteins belonging to the *VIM* family, many of which are PEGs (Hsieh *et al.* 2011, Belmonte *et al.* 2013). The roles of VIM genes in endosperm development are yet to be

examined but presumably resemble, to various extents, the effects of the loss of *MET1*. *MEE57* and several *VIM* genes are also imprinted, suggesting that the unique endosperm DNA methylation landscape may be relevant not only for regulating the transcription of imprinted genes but also for directing the function of proteins produced from imprinted genes. Other imprinted genes encode transcription factors whose binding is methylation sensitive. *DEL2,* a PEG, encodes a transcription factor that binds methylated DNA (O'Malley *et al.* 2016). It is therefore plausible that DNA methylation may regulate *DEL2*'s effects on gene expression. One of the challenges for the future will be deciphering the effects of DNA methylation on the functions of imprinted genes and vice versa.

Non-CG methylation

CHG methylation is maintained by CMT3 in *A. thaliana* (Stroud, *et al.* 2013). CMT3 orthologs are widely distributed in plants and have been identified in maize and rice (Lindroth 2001, Folsom *et al.* 2014, Bewick *et al.* 2016). CMT3 is expressed at relatively low levels in the endosperm (**Figure 2**) and has not been implicated in imprinting in *A. thaliana*. CMT3 might have roles in imprinting in other species, like *A. lyrata* or maize, where CHG methylation has been linked to imprinted genes (Zhang *et al.* 2014, Klosinska *et al.* 2016).

CHH methylation is added *de novo* either by CMT2 or DRM2 (Stroud, Do, *et al.* 2013). CMT2 is recruited to target sites by H3K9 methylation (Stroud, Do, *et al.* 2013). By reporter assays, CMT2 is not expressed in the central cell or the endosperm (Jullien *et al.* 2012) and no role in imprinting has been demonstrated for CMT2. DRM2 is recruited to target sites by small RNAs (RNA-directed DNA methylation or RdDM) (Matzke and Mosher 2014). The plant specific RNA Polymerase 4 (RNAP4) is one source of primary non-coding transcripts that are processed into small RNAs. DRM2 but not RNAP4 is expressed in the central cell (Vu *et al.* 2013), suggesting there may be little or no activity from the canonical RdDM pathway. One possibility is that a non-canonical RdDM pathway uses DRM2 to target DNA methylation (Panda *et al.* 2016). *RNAP4* expression levels are initially low in the endosperm (**Figure 2**), but increase at later stages of seed development, coincident with an increase in CHH methylation (Jullien *et al.* 2012, Belmonte *et al.* 2013). Several threads of evidence tie this small RNA pathway to imprinting. Small RNAs that map to imprinted genes have been identified in whole seed samples of *A. thaliana,* as well as in rice and maize endosperm (Rodrigues *et al.* 2013, Pignatta *et al.*

2014, Xin *et al.* 2014). For two loci, RNAP4 activity has also been shown to act in the paternal parent to establish a repressed state that silences the paternal allele of maternally expressed imprinted genes (MEGs) in the endosperm (Vu *et al.* 2013). At another maternally imprinted locus, *AtBMI1C,* imprinting is lost in mutants of the DICER ortholog *DCL4,* which functions downstream of RNAP4 to make small RNAs (Bratzel *et al.* 2012). A major challenge that remains in the field is to demonstrate the extent of the small RNA pathway's role in imprinting.

Sperm cell methylation

The sperm cell epigenome is also an important contributor to endosperm development, and loss of paternal DNA methylation impacts seed development and imprinting (Hsieh *et al.* 2011, Schatlowski *et al.* 2014). *A. thaliana* sperm cell transcriptomes and methylomes paint an interesting picture of the epigenetic machinery at work in these cells (Borges *et al.* 2008, Calarco *et al.* 2012). Expression of DNA demethylases is not detected (Schoft *et al.* 2011). Among the DNA methylation maintenance machinery, MET1 and the chromatin remodeler DDM1 are highly expressed while the CHG maintenance methyltransferase CMT3 is either absent or expressed at very low levels (Borges *et al.* 2008). The mRNA of *DRM2* could be detected but protein from a *DRM2* reporter transgene was undetectable in sperm cells (Borges *et al.* 2008, Calarco *et al.* 2012). Expression of several members of the canonical RdDM pathway, including RNAP4 sub-units and *DCL3* and *AGO4*, is absent in sperm cells (Borges *et al.* 2008). Instead, *DCL1, AGO5, AGO6, AGO9* are expressed in sperm, and it has been argued that this expression pattern presages a novel small RNA pathways in sperm cells (Borges *et al.* 2008). Consistent with this expressed machinery, sperm genomes show extensive CG methylation but very little CHH methylation (Calarco *et al.* 2012). Intriguingly, comparison of CG and CHH methylation at a limited number of imprinted genes showed that maternally expressed genes, compared to paternally expressed genes, are more methylated and are targeted by more small RNA (Calarco *et al.* 2012). However, the absence of the canonical RdDM pathway strongly suggests that either one of the non-canonical pathways or an yet undiscovered small RNA pathway targets DNA methylation in sperm (Matzke *et al.* 2015). It has also been proposed that small RNA produced in the vegetative cell (a non-gametic cell that encompasses the sperm cells in the pollen) influence DNA methylation in the sperm cell (Slotkin *et al.* 2009, Martínez *et al.* 2016). The functional relevance of such a process remains unclear.

DNA METHYLATION and IMPRINTING

DNA methylation, mostly studied in non-seed tissues, has been associated with regulation of transcription in *cis* as well as more global effects. *Cis* regulatory effects may be brought about by a multitude of mechanisms. Studies in somatic tissues as well as biochemical observations have shown that DNA methylation may promote or inhibit transcription factor recruitment (O'Malley *et al.* 2016). DNA methylation also modulates histone methylation. At the global level, DNA methylation may play roles in regulating chromatin organization (Soppe 2002) and accessibility (Shu *et al.* 2012) as well as in pairing between homologous chromosomal loci (Watanabe *et al.* 2005).

Single gene and whole genome DNA methylation studies have revealed an association between DNA methylation and imprinted genes. Broadly, the methylation pattern at imprinted genes mirrors the maternal hypomethylation seen at other genes in the genome. In *A. thaliana*, genome-wide studies find that 40% of MEGs and 62% of PEGs are associated with endosperm-specific CG DNA hypomethylation (Pignatta *et al.* 2014). This pattern is also seen in other species. In *A. lyrata*, half of all MEGs and a third of all PEGs are associated with CG hypomethylation on the maternal alleles (Klosinska *et al.* 2016). In maize, CG methylation is reduced at the maternal alleles of 54% of MEGs and 60% of PEGs (Zhang *et al.* 2014). Similarly, in rice, both MEGs and PEGs are also associated with DMRs (Rodrigues *et al.* 2013). A survey of three MEGs and four PEGs in Sorghum found that three PEGs were associated with CG hypomethylation in the endosperm (Zhang *et al.* 2016). It is likely that the number of imprinted genes associated with DMRs is under-estimated because of the limitations of analyses carried out thus far. Most studies have fixed parameters as to the level of hypomethylation at DMRs that is considered biologically relevant, as well as the distance of the DMR from the imprinted gene. DMRs that are farther away or less differentially methylated are unlikely to be identified as being linked to an imprinted gene. In summary, all of the studies show a strong association between maternal allele hypomethylation and imprinted expression. This correlation is also supported by examples of "allelic" or "strain-specific" imprinting. Allele-specific expression analyses among multiple strains of *Arabidopsis* identified 12 imprinted genes that are imprinted in some strains but not others (Pignatta *et al.* 2014). Interestingly, 10 of these genes

were associated with sites of inter-strain DNA methylation variation, arguing a significant role for DNA methylation in regulating imprinted expression (Pignatta *et al.* 2014).

DNA methylation and MEGs

In the endosperm, a seemingly paradoxical observation is that the loss of maternal allele DNA methylation is associated with both maternally expressed imprinted genes and paternally expressed imprinted genes. How then does the loss of DNA methylation activate the transcription of maternal alleles of MEGs but repress the maternal alleles of PEGs? The role of DNA methylation at several assayed MEGs is straightforward – it is repressive. In *A. thaliana,* loss of maternal *DME,* which demethylates maternal alleles, leads to reduced expression of nine MEGs (Hsieh *et al.* 2011). Correspondingly, in these nine cases, the loss of paternal *MET1* and hence a loss of DNA methylation on paternal alleles, leads to derepression of paternal alleles (Hsieh *et al.* 2011). In-depth study of individual loci also supports this view. *FWA* is repressed throughout *A. thaliana* development except in the endosperm, where it is expressed from the maternal allele (Kinoshita 2004). The upstream region of *FWA* has repeats that are maternally hypomethylated in the endosperm (Kinoshita 2004). The paternal allele repeats remain highly methylated, consistent with methylation and repression of *FWA* in somatic tissues. These data suggest that loss of DNA methylation relieves repression of the maternal alleles (**Figure 3**). Consistent with this view, inheritance of mutations in *MET1* from the paternal parent causes increased expression of the *FWA* paternal allele in endosperm and, conversely, loss of *DME* in the maternal parent leads to loss of maternal allele expression (Kinoshita 2004).

Single gene studies have also helped identify how the CG maintenance methylation pathway and the RdDM pathway act together to imprint a single locus. *SDC* and *MOP9.*5 are MEGs (Bratzel *et al.* 2012, Vu *et al.* 2013). The well-studied *SDC* locus produces a transcript that is initiated 78 bp downstream of seven 32 bp tandem repeats (Henderson and Jacobsen 2008). In somatic tissues, the repeats are targeted for DNA methylation by RdDM and CMT3 methylation pathways (Henderson and Jacobsen 2008). This non-CG methylation lays a foundation that allows *MET1* to expand CG methylation into regions upstream and downstream of the repeats (Henderson and Jacobsen 2008). Low levels of RdDM and DNA methyltransferases along with the activity of *DME* in the central cell presumably reduces DNA methylation over the upstream repeats and transcriptional start site (TSS) of *SDC* and allows for activation of the maternal allele

(Vu *et al.* 2013). On the other hand, the paternal allele remains methylated and repressed. This methylation is likely inherited from the male parent's diploid tissues through gametogenesis and is retained in the endosperm (Vu *et al.* 2013). The mechanism of *MOP9.5* imprinting is a little different (Vu *et al.* 2013). RdDM generates the repressive imprint on *MOP9.5*'s paternal allele in the paternal parent's diploid tissues prior to gametogenesis. However, the activation of the maternal allele is not dependent on *DME*. Instead it may be activated either by one of the other 5-methylcytosine DNA glycosylases or simply by the absence of RdDM in the central cell.

Examination of individual loci in rice also supports a repressive role for DNA methylation at maternal alleles of MEGs, which is relieved in the endosperm (Du *et al.* 2014). The maternal alleles of *Os01g69110* and *Os11g36470* produce endosperm-specific transcripts. The endosperm-specific transcriptional start sites are closely associated with DNA hypomethylation (**Figure 3**). Additionally, the abundance of these transcripts increases in seedlings (where they are usually absent) grown on chemicals that induce DNA hypomethylation.

Detailed analyses of individual imprinted genes also highlight the potential for DNA methylation to play roles in more complex regulatory circuits at imprinted genes. In *Arabidopsis,* the paternal and maternal alleles of *MEDEA* (*MEA*) are regulated by different mechanisms (Gehring *et al.* 2006). *MEA,* which encodes a PRC2 subunit, is normally expressed from the maternal allele. Transcription of the maternal allele is initiated in the central cell and is the result of the *DME*-dependent demethylation of a repetitive sequence downstream of *MEA.* However, the paternal *MEA* allele, which is normally methylated, is silenced in the endosperm not by DNA methylation, but by Polycomb mediated H3K27me3 (Gehring *et al.* 2006).

DNA methylation and PEGs

For PEGs, the maternal hypomethylated allele is the silent allele. Why is DNA hypomethylation correlated with transcriptional repression for these genes? The most popular model envisages an antagonistic relationship between DNA methylation and PRC2 mediated H3K27 methylation, which have been suggested to be mutually exclusive marks (**Figure 3**). This hypothesis is supported by genome-wide maps of H3K27me3 and allele-specific expression data from mutants with compromised PRC2 activity (Hsieh *et al.* 2011, Moreno-Romero *et al.* 2016). In *A. thaliana*, maternal alleles of 35 out of 43 PEGs were preferentially marked by H3K27me3 and loss of the Polycomb component *FIE* led to a loss of imprinting at these PEGs (Hsieh *et al.* 2011,

Moreno-Romero *et al.* 2016). Additionally, the maternal alleles of 93 paternally biased genes were also marked by H3K27me3 (Moreno-Romero *et al.* 2016). In maize, 36 out of 42 assayed PEGs had preferential H3K27me3 on the maternal alleles (Zhang *et al.* 2014). At the individual gene level, *OsYUCCA11* is expressed from a methylated paternal allele in the rice endosperm. The maternal allele has little or no CG methylation in the upstream region of the gene and gene body and is preferentially marked by H3K27me3.

In reality, the above model is probably more complicated; studies in maize and *Arabidopsis* show that DNA methylation and H3K27 methylation can occur together (Zhang *et al.* 2014, Moreno-Romero *et al.* 2016). While it remains possible that some of the overlap observed in genomic studies stems from different endosperm cell types having different epigenetic states, there are mechanisms for their co-occurrence. In vitro pull downs and yeast two hybrid assays show that the Arabidopsis DNA methyltransferase MET1 can, like its mammalian counterpart, physically interact with the PRC2 histone lysine methyltransferase subunit MEDEA; it has been suggested that this interaction plays role in imprinting at two different loci (Viré *et al.* 2005, Schmidt *et al.* 2013). In rice, a H3K27 methyltransferase, SGD711, physically interacts with DRM2 and a methylated CHG binding protein, SGD703, and targets histone methylation to regions with high CHH and CHG methylation (Zhou, Liu, *et al.* 2016).

Interestingly, in *A. lyrata* the relationship between PEGs and DNA methylation is somewhat more complicated than in *A. thaliana*. Like in *A. thaliana*, PEGs are associated with maternal allele DNA hypomethylation in gene-flanking regions. However, DNA methylation possibly has an additional repressive role on the maternal alleles of many PEGs, which are associated with increased CHG methylation over the gene body (Klosinska *et al.* 2016). An analysis of all genes, imprinted and non-imprinted, showed that the level of gain of gene body CHG methylation correlated with decreased maternal allele contribution (Klosinska *et al.* 2016). One possibility is that CHG methylation substitutes for H3K27me3 as a repressive mark on the maternal allele of *A. lyrata* PEGs, although the H3K27me3 landscape in *A. lyrata* endosperm is presently unknown.

Location matters

In the endosperm, at least part of the relationship between DNA methylation and imprinted gene expression is likely dependent on where the methylation in location in relation to gene features

and to Polycomb response elements. The hypomethylation on the maternal alleles of maize PEGs includes both the upstream region of the gene as well as the gene body, whereas the hypomethylation at MEGs is associated only with the gene body (Zhang *et al.* 2014). In rice, PEGs are often associated with embryo-endosperm DMRs in gene bodies and MEGs are associated with DMRs over the transcriptional start site and the 3' end of the gene (Rodrigues *et al.* 2013) While the effects of hypomethylation of the TSS can be predicted to impact transcription, the effects of DMRs that stretch over other parts of the gene or the gene's neighborhood remain unclear. Some of these DMRs uncover cryptic transcriptional start sites and promote the transcription of longer or truncated transcripts (**Figure 3**) (Du *et al.* 2014). In *A. lyrata,* the CG DMRs on gene bodies have been proposed to allow the establishment of repressive CHG methylation (Klosinska *et al.* 2016). DMRs may also help modulate the availability of binding sites for proteins that recruit Polycomb complexes. The mapping of H3K27me3 sites in the endosperm genome, the availability of DNA methylomes, and recent advances in identifying Polycomb response elements in plants will help uncover some of these mechanisms (Deng *et al.* 2013, Zhang *et al.* 2014, Hecker *et al.* 2015, Moreno-Romero *et al.* 2016, Zhou, Hartwig, *et al.* 2016).

THE FUTURE

In the past, studies exploring the endosperm required tedious, painstaking dissection, making it difficult to obtain the large amounts of material needed for techniques like ChIP. However, more recently, the declining requirement of material for whole genome approaches, as well as the use of techniques such as FACS and INTACT to isolate gamete and endosperm nuclei are opening up the endosperm and imprinting field (Moreno-Romero *et al.* 2016, Park *et al.* 2016). These techniques will help increase experimental throughput and enable the rapid examination of the epigenomes and imprinting in a number of mutants. Another development that will have a major impact on the field is the explosion in the number of non-model systems whose endosperm has come under the gaze of molecular biology and genomics (Hatorangan *et al.* 2016, Klosinska *et al.* 2016, Oneal *et al.* 2016, Xu *et al.* 2016, Zhang *et al.* 2016). Evolution promises that innumerable variations in imprinting molecular mechanisms lie beyond the laboratory models. Studies in diverse species will help us further understand endosperm methylation dynamics,

imprinting function, the evolution of imprinting, and the role of DNA methylation in imprinted expression.

Declarations of Interest

The authors report no declarations of interest. Research on imprinting in the authors' lab is supported by NSF grant 1453459.

FIGURE LEGENDS

Figure 1: Fertilization, seed development, and the origin DNA methylation differences.

The ovule envelops the female gametophyte, which contains two gametes, a haploid egg cell and a diploid central cell. The 5-methylcytosine DNA glycosylase DME is expressed in the central cell. The pollen grain consists of two haploid sperm cells embedded in a vegetative cell. The fusion of a haploid sperm cell and the haploid egg cell forms the diploid embryo. The fusion of a haploid sperm cell and the central cell forms the triploid endosperm. The seed coat is diploid and derived from the ovule and is thus completely maternal tissue. In the endosperm, maternal genomes are hypomethylated compared to the paternal genome. Embryo maternal and paternal genomes are not differentially methylated. This reflects differences in *DME* and *MET1* expression in the gametes before fertilization. Blue bars, DNA. Black lollipops, methylcytosine. Clear lollipops, unmethylated cytosine.

Figure 2: Developmental expression of key DNA methyltransferases and RdDM components in endosperm compartments and embryo.

A) *MET1* is expressed more highly in the embryo than in any part of the endosperm. B) *MEE57* expression is almost exclusive to the chalazal endosperm. C) *MET2A* and *MET2B* (AT4G08990 and AT4G14140) are expressed in the endosperm but are almost absent from the embryo. D) *CMT3* is more highly expressed in the embryo. E) *NRPD1A* (a subunit of RNA Polymerase 4) is expressed in some regions of the endosperm. F) *AGO4* expression increases in the endosperm during development; in the embryo it is always highly expressed. G) *RDR2* expression increases during development. H) *DRM2* expression increases in the endosperm during development. Data was compiled from Belmonte et al (Belmonte *et al.* 2013); see reference for experimental details. The Y-Axis represents RMA averaged signal. A bar bordered with a solid black line indicates that the gene is considered to be present or expressed in all replicate microarrays. A broken black line indicates that the gene was absent or not expressed in some replicate microarrays. A yellow border indicates that the gene is absent, expressed at very low levels or not expressed in all microarray replicates. Data for each tissue type are arranged by advancing developmental stage from left to right.

Figure 3: Different models on the relationship between maternal allele hypomethylation and gene imprinting. Hypomethylation affects transcription in different ways at different loci. A) Loss of DNA methylation relieves repression of maternal alleles at some loci. Paternal alleles remain methylated and repressed. B) DNA demethylation may uncover cryptic transcriptional start sites. Paternal alleles are transcribed from the typical start site. C) Maternal allele hypomethylation may allow recruitment of PRC2, leading to methylation of histone H3 on lysine 27 (H3K27me3). The paternal allele remains relatively more methylated and thus refractory to H3K27me3 modification. Black lollipops represent methylated cytosines and clear lollipops represent unmethylated cytosines. Triangles represent H3K27me3. Lines represent upstream regions and boxes represent genes. Wavy lines represent transcripts. Bold right angled arrow represents an active transcriptional start site. Faded right angled arrow represents a repressed transcriptional start site.

References

1. Achour, M., Jacq, X., Rondé,et al. (2008). The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene*, 27 (15), 2187–97.
2. Andreuzza, S., Li, J., Guitton, A.-E., et al. (2010). DNA LIGASE I exerts a maternal effect on seed development in Arabidopsis thaliana. *Development (Cambridge, England)*, 137 (1), 73–81.
3. Belmonte, M.F., Kirkbride, R.C., Stone, S.L., et al. (2013). Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *Proceedings of the National Academy of Sciences*, 110 (5), E435–E444.
4. Bewick, A.J., Ji, L., Niederhuth, C.E., et al. (2016). On the origin and evolutionary consequences of gene body DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America*, 113 (32), 9111–6.
5. Borges, F., Gomes, G., Gardner, R., et al. (2008). Comparative Transcriptomics of Arabidopsis Sperm Cells. *PLANT PHYSIOLOGY*, 148 (2), 1168–1181.
6. Bratzel, F., Yang, C., Angelova, A., et al. (2012). Regulation of the New Arabidopsis Imprinted Gene AtBMI1C Requires the Interplay

of Different Epigenetic Mechanisms. *Molecular Plant*, 5 (1), 260–269.

7. Buzas, D.M., Nakamura, M., and Kinoshita, T. (2014). Epigenetic role for the conserved Fe-S cluster biogenesis protein AtDRE2 in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 111 (37), 13565–13570.
8. Calarco, J.P., Borges, F., Donoghue, M.T.A., Van Ex, F., et al. (2012). Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA. *Cell*, 151 (1), 194–205.
9. Cao, X. and Jacobsen, S.E. (2002). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proceedings of the National Academy of Sciences*, 99 (Supplement 4), 16491–16498.
10. Choi, Y., Gehring, M., Johnson, L., et al. (2002). DEMETER, a DNA Glycosylase Domain Protein, Is Required for Endosperm Gene Imprinting and Seed Viability in Arabidopsis. *Cell*, 110 (1), 33–42.
11. Costa, L.M., Yuan, J., Rouster, J., et al. (2012). Maternal control of nutrient allocation in plant seeds by genomic imprinting. *Current biology : CB*, 22 (2), 160–5.
12. Deng, W., Buzas, D.M., Ying, H., et al. (2013). Arabidopsis Polycomb Repressive Complex 2 binding sites contain putative GAGA factor binding motifs within coding regions of genes. *BMC genomics*, 14, 593.
13. Du, M., Luo, M., Zhang, R., et al. (2014). Imprinting in rice: the role of DNA and histone methylation in modulating parent-of-origin specific expression and determining transcript start sites. *The Plant Journal*, 79 (2), 232–242.
14. Figueiredo, D.D., Batista, R.A., Roszak, P.J., and Köhler, C., 2015. Auxin production couples endosperm development to fertilization. *Nature Plants*, 1 (12), 15184.
15. Folsom, J.J., Begcy, K., Hao, X., et al. (2014). Rice Fertilization-Independent Endosperm1 Regulates Seed Size under Heat Stress by Controlling Early Endosperm Development. *PLANT PHYSIOLOGY*, 165 (1), 238–248.
16. Gehring, M., Bubb, K.L., and Henikoff, S.(2009). Extensive Demethylation of Repetitive Elements During Seed Development Underlies Gene Imprinting. *Science*, 324 (5933), 1447–1451.
17. Gehring, M., Huh, J.H., Hsieh, T.-F., et al. (2006). DEMETER

DNA Glycosylase Establishes MEDEA Polycomb Gene Self-Imprinting by Allele-Specific Demethylation. *Cell*, 124 (3), 495–506.

18. Gehring, M., Missirian, V., and Henikoff, S., 2011. Genomic analysis of parent-of-origin allelic expression in arabidopsis thaliana seeds. *PLoS ONE*, 6 (8).
19. Gent, J.I., Madzima, T.F., Bader, R., et al. (2014). Accessible DNA and Relative Depletion of H3K9me2 at Maize Loci Undergoing RNA-Directed DNA Methylation. *The Plant Cell Online*, 26 (12), 4903–4917.
20. Haig, D., 2013. Kin conflict in seed development: an interdependent but fractious collective. *Annual review of cell and developmental biology*, 29, 189–211.
21. Hatorangan, M.R., Laenen, B., Steige, et al. (2016). Rapid Evolution of Genomic Imprinting in Two Species of the Brassicaceae. *The Plant cell*.
22. Hecker, A., Brand, L.H., Peter, S., et al. (2015). The Arabidopsis GAGA-Binding Factor BASIC PENTACYSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs. *Plant physiology*, 168 (3), 1013–24.
23. Henderson, I.R. and Jacobsen, S.E., 2008. Tandem repeats upstream of the Arabidopsis endogene SDC recruit non-CG DNA methylation and initiate siRNA spreading. *Genes & Development*, 22 (12), 1597–1606.
24. Hsieh, T.-F., Ibarra, C.A., Silva, P., et al. (2009). Genome-Wide Demethylation of Arabidopsis Endosperm. *Science*, 324 (5933), 1451–1454.
25. Hsieh, T.F., Shin, J., Uzawa, R., et al. (2011). Regulation of imprinted gene expression in Arabidopsis endosperm. *Proceedings of the National Academy of Sciences*, 108 (5), 1755–1762.
26. Hu, L., Li, N., Xu, C., et al. (2014). Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proceedings of the National Academy of Sciences*, 111 (29), 10642–10647.
27. Ibarra, C.A., Feng, X., Schoft, V.K., et al. (2012). Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes. *Science*, 337 (6100), 1360–1364.

28. Ikeda, Y., Kinoshita, Y., Susaki, D., et al. (2011). HMG Domain Containing SSRP1 Is Required for DNA Demethylation and Genomic Imprinting in Arabidopsis. *Developmental Cell*, 21 (3), 589–596.
29. Johnston, A.J., Matveeva, E., Kirioukhova, O., et al. (2008). A Dynamic Reciprocal RBR-PRC2 Regulatory Circuit Controls Arabidopsis Gametophyte Development. *Current Biology*, 18 (21), 1680–1686.
30. Jullien, P.E., Kinoshita, T., Ohad, N., and Berger, F. (2006). Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. *The Plant cell*, 18 (6), 1360–72.
31. Jullien, P.E., Mosquna, A., Ingouff, M., et al. (2008). Retinoblastoma and Its Binding Partner MSI1 Control Imprinting in Arabidopsis. *PLoS Biology*, 6 (8), e194.
32. Jullien, P.E., Susaki, D., Yelagandula, R., et al. (2012). DNA Methylation Dynamics during Sexual Reproduction in Arabidopsis thaliana. *Current Biology*, 22 (19), 1825–1830.
33. Kapazoglou, A., Drosou, V., Argiriou, A., and Tsaftaris, A.S.(2013). The study of a barley epigenetic regulator, HvDME, in seed development and under drought. *BMC Plant Biology*, 13 (1), 172.
34. Kinoshita, T. (2004). One-Way Control of FWA Imprinting in Arabidopsis Endosperm by DNA Methylation. *Science*, 303 (5657), 521–523.
35. Kiyosue, T., Ohad, N., Yadegari, R., et al. (1999). Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 96 (7), 4186–91.
36. Klosinska, M., Picard, C.L., and Gehring, M.(2016). Conserved imprinting associated with unique epigenetic signatures in the Arabidopsis genus. *Nature Plants*, 2, 16145.
37. Law, J.A. and Jacobsen, S.E.(2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, 11 (3), 204–220.
38. Li, J. and Berger, F.(2012). Endosperm: Food for humankind and fodder for scientific discoveries. *New Phytologist*, 195 (2), 290–305.

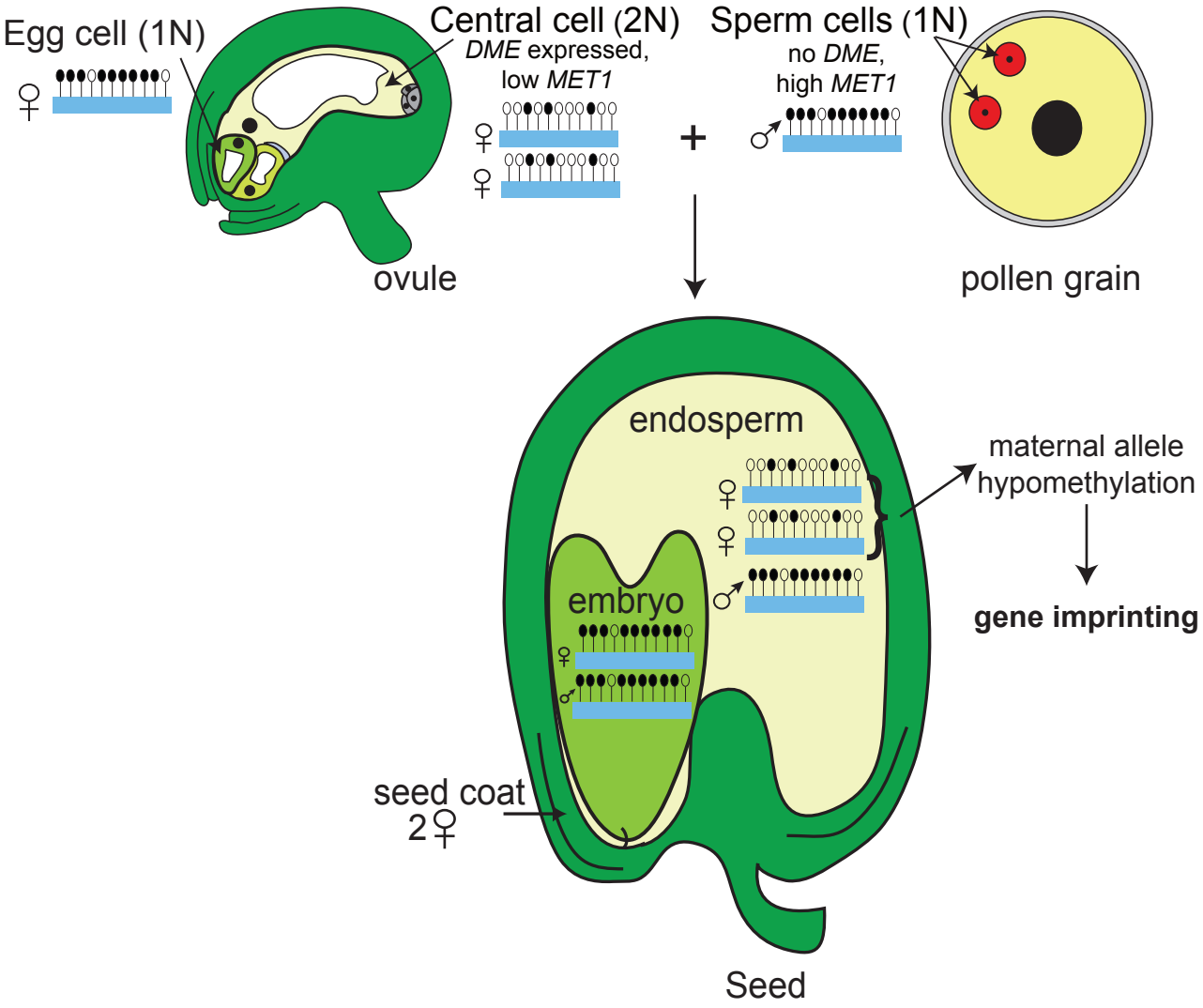
39. Li, Y., Córdoba-Cañero, D., Qian, W., et al. (2015). An AP Endonuclease Functions in Active DNA Demethylation and Gene Imprinting in Arabidopsis. *PLoS Genetics*, 11 (1), e1004905.
40. Lindroth, A.M. (2001). Requirement of CHROMOMETHYLASE3 for Maintenance of CpXpG Methylation. *Science*, 292 (5524), 2077–2080.
41. Luo, M., Taylor, J.M., Spriggs, et al. (2011). A Genome-Wide Survey of Imprinted Genes in Rice Seeds Reveals Imprinting Primarily Occurs in the Endosperm. *PLoS Genetics*, 7 (6), e1002125.
42. Martínez, G., Panda, K., Köhler, C., and Slotkin, R.K.(2016). Silencing in sperm cells is directed by RNA movement from the surrounding nurse cell. *Nature Plants*, 2 (4), 16030.
43. Maruyama, D., Völz, R., Takeuchi, H., et al. (2015). Rapid Elimination of the Persistent Synergid through a Cell Fusion Mechanism. *Cell*, 161 (4), 907–918.
44. Matzke, M.A., Kanno, T., and Matzke, A.J.M.(2015). RNA-Directed DNA Methylation: The Evolution of a Complex Epigenetic Pathway in Flowering Plants. *Annual Review of Plant Biology*, 66 (1), 243–267.
45. Matzke, M.A. and Mosher, R.A., (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics*, 15 (6), 394–408.
46. Mok, Y.G., Uzawa, R., Lee, J., et al. (2010). Domain structure of the DEMETER 5-methylcytosine DNA glycosylase. *Proceedings of the National Academy of Sciences*, 107 (45), 19225–19230.
47. Moreno-Romero, J., Jiang, H., Santos-González, J., and Köhler, C. (2016). Parental epigenetic asymmetry of PRC2-mediated histone modifications in the Arabidopsis endosperm. *The EMBO Journal*, 35 (12), 1298–1311.
48. Nakamura, M., Buzas, D.M., Kato, A., et al. (2013). The role of Arabidopsis thaliana NAR1, a cytosolic iron-sulfur cluster assembly component, in gametophytic gene expression and oxidative stress responses in vegetative tissue. *New Phytologist*, 199 (4), 925–935.
49. Nishiyama, A., Yamaguchi, L., Sharif, J., et al. (2013). Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. *Nature*, 502 (7470), 249–53.
50. O'Malley, R.C., Huang, S.C., Song, L., et al. (2016). Cistrome and

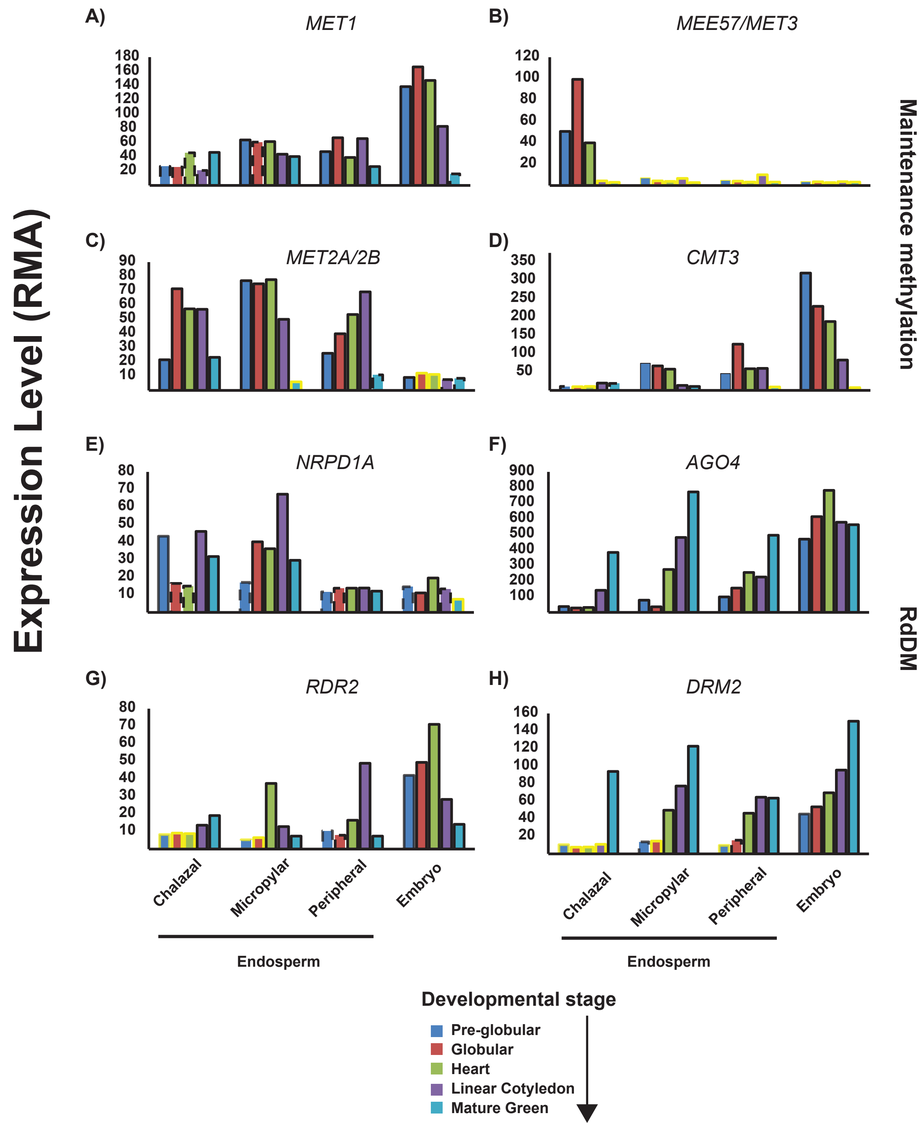
Epicistrome Features Shape the Regulatory DNA Landscape. *Cell*, 165 (5), 1280–1292.
51. Oneal, E., Willis, J.H., and Franks, R.G. (2016). Disruption of endosperm development is a major cause of hybrid seed inviability between Mimulus guttatus and Mimulus nudatus. *New Phytologist*, 210 (3), 1107–1120.
52. Ono, A., Yamaguchi, K., Fukada-Tanaka, S., et al. (2012). A null mutation of ROS1a for DNA demethylation in rice is not transmittable to progeny. *The Plant Journal*, 71 (4), 564–574.
53. Panda, K., Ji, L., Neumann, D.A., et al. (2016). Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biology*, 17 (1), 170.
54. Park, K., Kim, M.Y., Vickers, M., et al. (2016). DNA demethylation is initiated in the central cells of Arabidopsis and rice. *Proceedings of the National Academy of Sciences*, 201619047.
55. Pignatta, D., Erdmann, R.M., Scheer, E., et al. (2014). Natural epigenetic polymorphisms lead to intraspecific variation in Arabidopsis gene imprinting. *eLife*, 3.
56. Pillot, M., Baroux, C., Vazquez, M.A., et al. (2010). Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in Arabidopsis. *The Plant cell*, 22 (2), 307–320.
57. Quadrana, L., Bortolini Silveira, A., Mayhew, G.F., et al. (2016). The Arabidopsis thaliana mobilome and its impact at the species level. *eLife*, 5.
58. Rea, M., Zheng, W., Chen, M., et al. (2012). Histone H1 affects gene imprinting and DNA methylation in Arabidopsis. *The Plant journal : for cell and molecular biology*, 71 (5), 776–86.
59. Rodrigues, J.A., Ruan, R., Nishimura, T., et al. (2013). Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. *Proceedings of the National Academy of Sciences of the United States of America*, 110 (19), 7934–9.
60. Ronemus, M.J., Galbiati, M., Ticknor, C., et al. (1996). Demethylation-Induced Developmental Pleiotropy in Arabidopsis. *Science*, 273 (5275), 654–657.
61. Schatlowski, N., Wolff, P., Santos-González, J., et al. (2014).

Hypomethylated pollen bypasses the interploidy hybridization barrier in Arabidopsis. *The Plant cell*, 26 (9), 3556–68.
62. Schmidt, A., Wöhrmann, H.J.P., Raissig, M.T., et al. (2013). The Polycomb group protein MEDEA and the DNA methyltransferase MET1 interact to repress autonomous endosperm development in Arabidopsis. *The Plant Journal*, 73 (5), 776–787.
63. Schoft, V.K., Chumak, N., Choi, Y., et al. (2011). Function of the DEMETER DNA glycosylase in the Arabidopsis thaliana male gametophyte. *Proc Natl Acad Sci USA*, 108 (19), 8042–7.
64. Schoft, V.K., Chumak, N., Mosiolek, M., et al. (2009). Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. *EMBO reports*, 10 (9), 1015–21.
65. Shook, M.S. and Richards, E.J.(2014). VIM proteins regulate transcription exclusively through the MET1 cytosine methylation pathway. *Epigenetics*, 9 (7), 980–986.
66. Shu, H., Wildhaber, T., Siretskiy, A., et al. (2012). Distinct modes of DNA accessibility in plant chromatin. *Nature Communications*, 3, 1281.
67. Slotkin, R.K., Vaughn, M., Borges, F., et al. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell*, 136 (3), 461–72.
68. Soppe, W.J.J., 2002. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *The EMBO Journal*, 21 (23), 6549–6559.
69. Spencer, H.G. and Clark, A.G., 2014. Non-conflict theories for the evolution of genomic imprinting. *Heredity*, 113 (2), 112–118.
70. Stroud, H., Do, T., Du, J., et al. (2013). Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nature Structural & Molecular Biology*, 21 (1), 64–72.
71. Stroud, H., Greenberg, M.V.C., Feng, S., et al. (2013). Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell*, 152 (1-2), 352–364.
72. Viré, E., Brenner, C., Deplus, R., et al. (2005). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*, 439 (7078), 871–874.
73. Vu, T.M., Nakamura, M., Calarco, J.P., et al. (2013). RNA-directed DNA methylation regulates parental genomic imprinting at several loci in Arabidopsis. *Development (Cambridge, England)*,

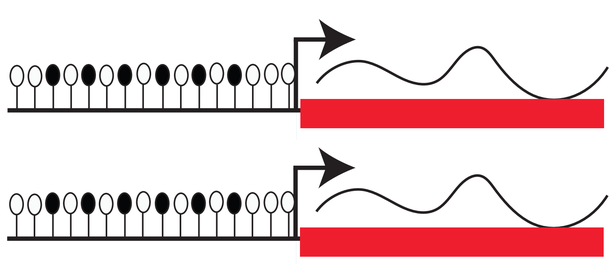
140 (14), 2953–60.
74. Wang, P., Xia, H., Zhang, Y., et al. (2015). Genome-wide high-resolution mapping of DNA methylation identifies epigenetic variation across embryo and endosperm in Maize (Zea may). *BMC Genomics*, 16 (1), 21.
75. Watanabe, K., Pecinka, A., Meister, A., et al. (2005). DNA hypomethylation reduces homologous pairing of inserted tandem repeat arrays in somatic nuclei of Arabidopsis thaliana. *The Plant Journal*, 44 (4), 531–540.
76. Waters, A.J., Bilinski, P., Eichten, S.R., et al. (2013). Comprehensive analysis of imprinted genes in maize reveals allelic variation for imprinting and limited conservation with other species. *Proceedings of the National Academy of Sciences*, 110 (48), 19639–19644.
77. Wolff, P., Weinhofer, I., Seguin, J., et al. (2011). High-Resolution analysis of parent-of-origin allelic expression in the Arabidopsis endosperm. *PLoS Genetics*, 7 (6).
78. Xiao, W., Gehring, M., Choi, Y., et al. (2003). Imprinting of the MEA Polycomb Gene Is Controlled by Antagonism between MET1 Methyltransferase and DME Glycosylase. *Developmental Cell*, 5 (6), 891–901.
79. Xin, M., Yang, R., Yao, Y., (2014). Dynamic parent-of-origin effects on small interfering RNA expression in the developing maize endosperm. *BMC Plant Biology*, 14 (1), 192.
80. Xing, M.-Q., Zhang, Y.-J., Zhou, S.-R., et al. (2015). Global Analysis Reveals the Crucial Roles of DNA Methylation during Rice Seed Development. *Plant physiology*, 168 (4), 1417–32.
81. Xu, W., Dai, M., Li, F., and Liu, A.(2014). Genomic imprinting, methylation and parent-of-origin effects in reciprocal hybrid endosperm of castor bean. *Nucleic Acids Research*, 42 (11), 6987–6998.
82. Xu, W., Yang, T., Dong, X., et al. (2016). Genomic DNA Methylation Analyses Reveal the Distinct Profiles in Castor Bean Seeds with Persistent Endosperms. *Plant physiology*, 171 (2), 1242–58.
83. Yamauchi, T., Johzuka-Hisatomi, Y., Terada, R., et al. (2014). The MET1b gene encoding a maintenance DNA methyltransferase is indispensable for normal development in rice. *Plant Molecular Biology*, 85 (3), 219–232.

84. Yelagandula, R., Stroud, H., Holec, S., et al. (2014). The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in arabidopsis. *Cell*, 158 (1), 98–109.
85. Zemach, A., Kim, M.Y., Silva, P., et al. (2010). Local DNA hypomethylation activates genes in rice endosperm. *Proceedings of the National Academy of Sciences of the United States of America*, 107 (43), 18729–18734.
86. Zhang, M., Li, N., He, W., et al. (2016). Genome-wide screen of genes imprinted in sorghum endosperm, and the roles of allelic differential cytosine methylation. *The Plant Journal*, 85 (3), 424–436.
87. Zhang, M., Xie, S., Dong, X., et al. (2014). Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize. *Genome Research*, 24 (1), 167–176.
88. Zhang, M., Zhao, H., Xie, S., et al. (2011). Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. *Proceedings of the National Academy of Sciences*, 108 (50), 20042–20047.
89. Zhang, X. and Jacobsen, S.E., 2006. Genetic Analyses of DNA Methyltransferases in Arabidopsis thaliana. *Cold Spring Harbor Symposia on Quantitative Biology*, 71, 439–447.
90. Zhou, S., Liu, X., Zhou, C., et al. (2016). Cooperation between the H3K27me3 chromatin marker and non-CG methylation in epigenetic regulation. *Plant physiology*.
91. Zhou, Y., Hartwig, B., James, G.V., et al. (2016). Complementary Activities of TELOMERE REPEAT BINDING Proteins and Polycomb Group Complexes in Transcriptional Regulation of Target Genes. *The Plant cell*, 28 (1), 87–101.

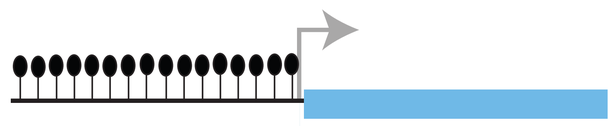

Egg cell (1N)
Central cell (2N)
DME expressed,
low MET1
Sperm cells (1N)
no DME,
high MET1
ovule
pollen grain
endosperm
embryo
maternal allele
hypomethylation
gene imprinting
seed coat
2♀
Seed


Expression Level (RMA)
A)
MET1
B)
MEE57/MET3
C)
MET2A/2B
D)
CMT3
E)
NRPD1A
F)
AGO4
G)
RDR2
H)
DRM2
Maintenance methylation
RdDM
Chalazal
Micropylar
Peripheral
Embryo
Endosperm
Developmental stage
Pre-globular
Globular
Heart
Linear Cotyledon
Mature Green

A) Hypomethylation activates transcription

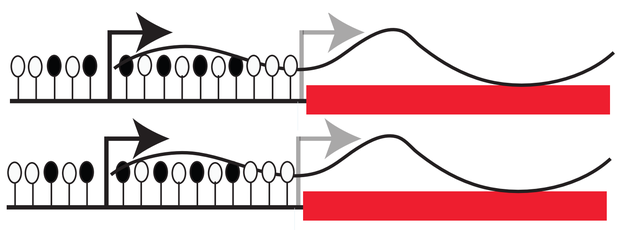


Maternal allele active due to loss of DNA methylation

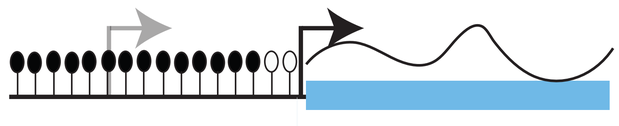


Paternal allele repressed by DNA methylation

B) Hypomethylation activates transcription from a cryptic upstream start site

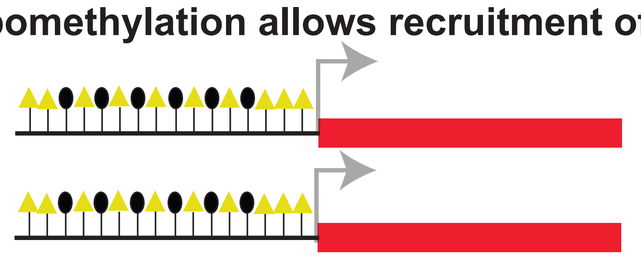


Loss of DNA methylation on maternal alleles uncovers upstream start site

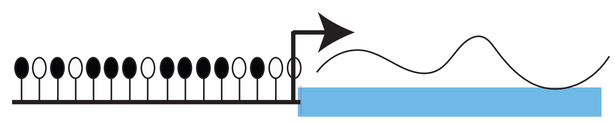


DNA methylation masks upstream start site

C) Hypomethylation allows recruitment of repressive H3K27Me3



Maternal alleles repressed by H3K27me3 deposited by PRC2 complexes



Paternal allele active because DNA methylation blocks H3K27me3