Minireview

Reprogramming the epigenome during germline and seed development

Mark A Johnson and Judith Bender

Address: Brown University, Department of Molecular Biology, Cell Biology, and Biochemistry, 185 Meeting St., Providence, RI 02912, USA.

Correspondence: Judith Bender. Email: judith_bender@brown.edu

Abstract
Gene silencing by DNA methylation and small RNAs is globally reconfigured during gametogenesis in *Arabidopsis*, affecting transposon activity, gene regulation and development.

Epigenomics, the comprehensive study of how a genome is marked by histone modifications and 5-methylcytosine DNA methylation to provide regulatory information, is now in its heyday. Epigenomics has the power to uncover overall trends in chromatin dynamics as well as specific regulatory loci during development and disease progression. The reference plant *Arabidopsis thaliana* has led the way in developing epigenomics technologies because of its compact genome decorated with histone modifications and DNA methylation that serve similar gene-activating or gene-repressing functions as in mammalian genomes. For example, in *Arabidopsis*, transposons and repetitive sequences are targeted for DNA methylation and transcriptional gene silencing through a small-RNA-guided mechanism [1] analogous to the specialized Piwi-interacting RNA (piRNA) pathway in the mammalian germline [2].

Recent studies published in *Science, Nature* and *Cell* have used *Arabidopsis* epigenomics to query patterns of DNA methylation, transposable element expression and small RNA accumulation in reproductive cells [3] and during seed development [4-6]. The striking finding is that cells accompanying both the sperm and the embryo undergo genome-wide changes in DNA methylation that activate small RNA production from transposons and other repeats. This finding raises the intriguing possibility that the supporting cells are used as a source of small RNAs that can move into the germline and reconfigure the epigenome.

Successful fertilization and early embryo development in flowering plants depend on cells that support, but do not contribute, to the germline. During egg development, the haploid product of meiosis undergoes three subsequent cell divisions that generate the female gametophyte: a haploid egg cell surrounded by supporting cells, including a diploid central cell [7]. Similarly, the male meiotic product undergoes subsequent cell divisions that give rise to the male gametophyte: two haploid sperm cells housed in the pollen cytoplasm in close association with the haploid pollen nucleus, called the vegetative nucleus [8]. During fertilization, an extension of the pollen grain grows through floral tissue to deliver two sperm to a female gametophyte. One sperm fertilizes the egg to produce the embryo (green, Figure 1); the other sperm fertilizes the central cell to produce the endosperm (yellow, Figure 1). The endosperm feeds the embryo as it develops, analogously to the mammalian placenta. The endosperm is essentially a twin of the embryo except that it contains an extra copy of the maternal genome. Together, the endosperm and the embryo are the major components of seeds. The mature *Arabidopsis* seed is only 300 μm in diameter, and the pollen grain is only 25 μm in diameter, so gathering reproductive cells from these tissues in quantities sufficient for biochemical analysis requires a heroic effort.

Previous genetic studies of seed development in *Arabidopsis* revealed a critical role for the DEMETER DNA demethylase, which is expressed specifically in the central cell [9]. A handful of DEMETER target genes were also characterized, most of which have a promoter element that is demethylated in the central cell to activate maternal-specific expression in the endosperm [10]. However, the *PHERES1* target gene has a 3' region that causes maternal-specific repression when DNA methylation is removed, perhaps by unblocking an insulator sequence [4,11]. The known targets raised the possibility that DEMETER might act more globally to erase DNA methylation on the two maternal chromosomes versus the single paternal chromosome that contribute to the endosperm genome.

Endosperm genomics
Gehring *et al.* [4] set out to find additional *Arabidopsis* genes that show parent-of-origin-specific (imprinted) expression in the endosperm. Their strategy was to identify regions of the genome that have reduced DNA methylation in the endosperm, compared with the embryo, as candidates for imprinted loci. They dissected endosperm and embryo tissues from developing seeds and immunoprecipitated methylated DNA from each sample with an anti-5-methylcytosine antibody. They then used
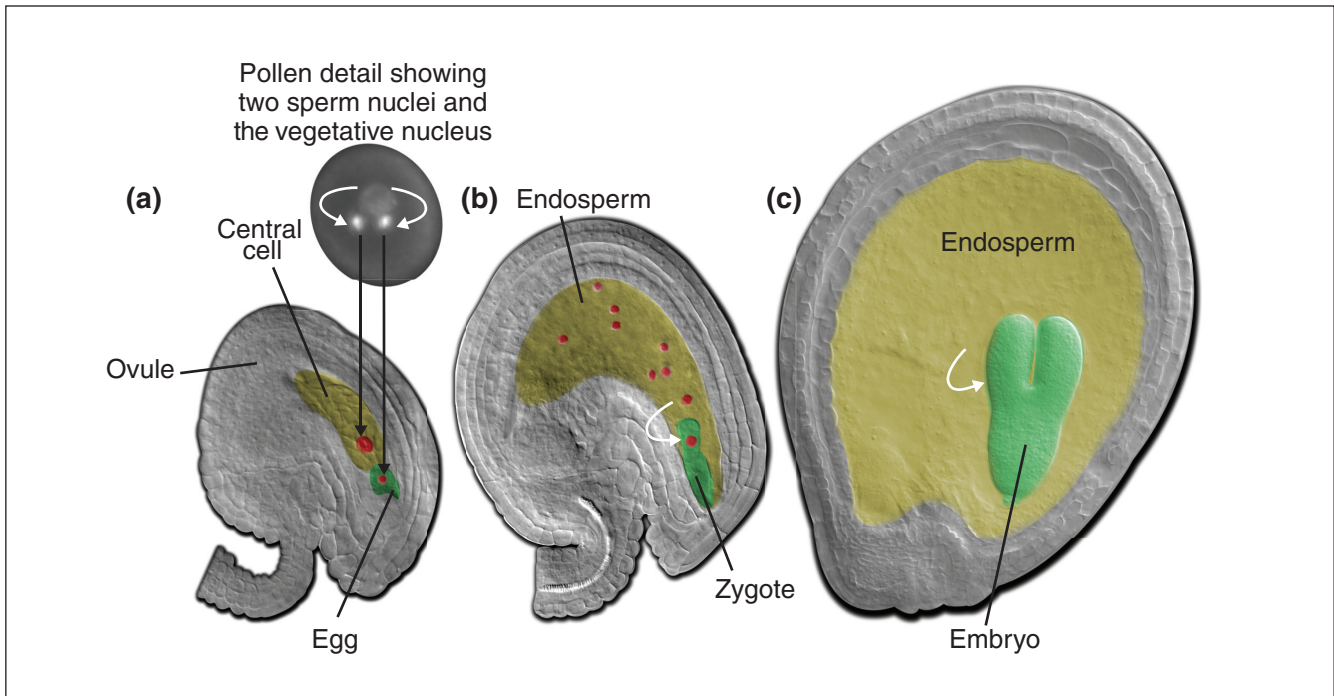
Figure 1

Flowering plant fertilization and embryo development. Micrographs of **(a)** pollen (detail of nuclei only) and ovule before fertilization, **(b)** ovule after fertilization, showing endosperm development (red dots are individual nuclei in the syncytial endosperm); and **(c)** a developing seed; all images are false-colored to indicate endosperm (yellow) and the embryo (green). All images are from *Arabidopsis*. In the pollen grain, the two sperm nuclei (fluorescent dots) are supported by the larger haploid vegetative nucleus; in the ovule, the haploid egg cell is supported by the diploid central cell. Angiosperms have a 'double fertilization', in which one sperm nucleus combines with the egg nucleus to generate the diploid zygote, and the other sperm nucleus combines with the larger central cell nucleus to generate a triploid endosperm. White arrows indicate the potential flow of small RNAs from non-inherited supporting cells into the gametes or the embryo.

high-throughput sequencing or hybridization to microarrays to determine which sequences were recovered. Analysis of the data revealed widespread reduced DNA methylation of transposons and other repeats that generate small RNAs in the endosperm.

This trend was partially reversed by mutation of *DEMETER*, indicating a role for active DNA demethylation. The authors then focused on genes that lay close to differentially methylated regions. They validated five new imprinted genes by showing allele-specific expression in endosperms generated by crosses between polymorphic parents. In each case, imprinted expression correlated with reduced DNA methylation on a transposon or repeated sequence within 1 kb of the coding sequence of the imprinted gene. On the basis of the characteristics of the validated genes, the authors estimate that approximately 50 genes in total will prove to have imprinted expression in the endosperm.

Hsieh *et al.* [6] used a different approach - shotgun bisulfite sequencing - to examine patterns of DNA methylation in embryo, endosperm and adult plant tissues.

Genomic DNA was treated with the mutagen sodium bisulfite to convert cytosines, but not 5-methylcytosines, to uracils. DNA was then subjected to high-throughput deep sequencing to visualize the global patterns of cytosine mutagenesis. The advantage of this approach over immunoprecipitating methylated DNA is that precise patterns of DNA methylation can be determined. This is an especially important consideration in plant genomes, where cytosines in different sequence contexts are methylated by different pathways. In particular, methylation of cytosines in asymmetric contexts is directed by small RNAs that correspond to the methylated sequence [1].

Genome-wide bisulfite sequencing of DNA from embryo, endosperm, and adult tissues showed dramatically different trends for methylation of cytosines in the symmetric context CG, which is the most prevalent context for *Arabidopsis* DNA methylation, compared with asymmetric contexts [6]. CG methylation was reduced in the endosperm in both gene and repeat sequences relative to the embryo, adult, and *demeter* mutant endosperm. This overall pattern is similar to that described by Gehring *et al.* [4]. However, Hsieh *et al.* [6] also found that asymmetric

cytosine methylation in repeat sequences was elevated in both endosperm and embryo relative to adult tissues, and that the *demeter* mutation strongly depleted asymmetric cytosine methylation in the endosperm.

A likely explanation for these context-specific methylation changes comes from a previous study showing that loss of CG methylation causes transcriptional reactivation of repeat sequences throughout the genome and a burst of new small RNAs, stimulating new methylation of asymmetric cytosines in the complementary regions [12]. Projecting this cascade onto the female gametophyte, DEMETER would first act genome-wide in the central cell to erase DNA methylation and activate synthesis of small RNAs. These small RNAs would trigger new asymmetric cytosine methylation on the two maternal and one paternal endosperm chromosomes, whereas demethylation in CG contexts on the maternal chromosomes would persist. In *demeter* mutants, the small RNAs would never be made, preventing methylation of asymmetric cytosines.

Consistent with this model, a recent small RNA high-throughput sequencing study by Mosher *et al.* [5] found that there is indeed a genome-wide burst of small RNAs from maternal genome transposons and repeats in the seed. A further test of this model will be to determine whether mutations that disrupt production of small RNAs block the increase in asymmetric cytosine methylation seen in endosperm compared to adult tissues.

Small RNA transfer to the germline
How then to account for the observed elevated levels of asymmetric cytosine methylation in the embryo [6]? The endosperm is known to transport nutrients to the embryo, so why not small RNAs? In this scenario, the embryo would maintain the DNA methylation patterns inherited from the egg and sperm genomes, but would reinforce them with an extra layer of asymmetric cytosine methylation mediated by imported small RNAs. The lockdown on repeat sequences would ensure stability of the plant genome during the critical early cell divisions. Small RNA import might be of limited extent, because Mosher *et al.* [5] found much lower levels of selected small RNAs in embryos versus endosperms using RNA blot analysis. It might nonetheless be possible to test the small RNA import model by determining whether asymmetric cytosine methylation is depleted in *demeter* mutant embryos as well as in *demeter* mutant endosperms.

The possibility of transport of small RNAs from a non-inherited supporting cell to the germline was previously suggested by analysis of transposon reactivation, small RNAs and DNA methylation in the *Arabidopsis* pollen and purified sperm cells [3]. Slotkin and colleagues [3] found that whole pollen grains (containing two sperm and one vegetative nucleus; Figure 1) showed several features diagnostic of

genome-wide transposon reactivation, including reduced DNA methylation on transposon sequences, high levels of transposon small RNAs, and new transposon insertions. However, purified sperm cells maintained full transposon DNA methylation. In addition, new transposon insertions were not transmitted to progeny plants. By process of elimination, the authors inferred that the loss of transposon DNA methylation, burst of small RNA synthesis, and transposon mobilization were occurring in the vegetative nucleus [3]. They proposed that the vegetative nucleus sacrifices its own genome integrity to supply protective small RNAs to sperm. In light of this proposal it will be interesting to see whether whole-genome bisulfite sequencing in purified sperm reveals the same patterns of elevated asymmetric cytosine methylation on repeats as observed in the embryo [6]. It will also be interesting to determine whether DEMETER or a related DNA demethylase is expressed specifically in the vegetative nucleus relative to the sperm nuclei.

Adaptive significance of DNA demethylation
Taken together, recent epigenomics studies agree that there is widespread reduced DNA methylation and production of transposon and repeat small RNAs in 'dead end' non-inherited cells that support the gametes and embryo. In the central cell of the female gametophyte, this sweeping change is achieved by DEMETER, which acts as a glycosylase to cleave 5-methylcytosines from the sugar-phosphate backbone [13]. The resulting abasic sites are replaced with unmethylated cytosines, presumably by base excision repair. When projected on to a genome-wide scale, this active DNA demethylation process has a high energetic cost and is probably mutagenic. There must, therefore, have been strong selective pressure that drove the evolution of this epigenetic remodeling strategy. Using the central cell and the vegetative nucleus as small RNA farms to feed the transposon-silencing needs of the egg, sperm and embryo could be sufficiently advantageous, given that mobile DNA transposons and retrotransposons are major components of plant genomes. In this view, genes with imprinted expression in the endosperm evolved by being caught up in the global DNA demethylation of the central cell nucleus by virtue of their proximity to DNA methylated repeats; genes that confer a growth advantage when activated or repressed from the maternal chromosomes were selected to maintain this mode of regulation [4].

Another idea is that the burst of endosperm small RNAs could be used to discriminate self-pollinated seeds from cross-pollinated hybrid seeds by *trans*-silencing non-self genes from a polymorphic paternal genome due to fortuitous sequence complementarity [5]. Hybrids with developmentally detrimental paternal gene silencing could be eliminated from the population early by seed abortion, while hybrids with developmentally beneficial paternal gene silencing would get an early edge in the competition for resources.

Although *Arabidopsis* shares the basics of reproductive development with other flowering plants (Figure 1), it has specialized features, including self-pollination and a small endosperm. In addition, only approximately 10% of the *Arabidopsis* genome is composed of transposon sequences [14]. It will therefore be important to analyze epigenomic patterns in reproductive tissues and seeds of other plants to assess the generality of DNA demethylation and small RNA activation. Clues already exist that DNA methylation is globally reduced in tobacco pollen [15] and in the maize endosperm [16]. An epigenomic analysis of maize reproduction will be especially interesting because maize is an out-crossing plant with a large endosperm that supports embryo development over a prolonged period. In addition, maize has a genome composed of more than 60% transposon-derived sequences [17], many of which are mobile [18]. Species-specific differences in endosperm epigenomics could ultimately reveal how seed size is determined, with implications for our future food and energy resources.

Acknowledgements

References

1. Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ: **RNA-mediated chromatin-based silencing in plants.** *Curr Opin Cell Biol* 2009, **21:**367-376.

2. Aravin AA, Sachidanandam R, Bourc'his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ: **A piRNA pathway primed by individual transposons is linked to *de novo* DNA methylation in mice.** *Mol Cell* 2008, **31:**785-799.

3. Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA, Martienssen RA: **Epigenetic reprogramming and small RNA silencing of transposable elements in pollen.** *Cell* 2009, **136:**461-472.

4. Gehring M, Bubb KL, Henikoff S: **Extensive demethylation of repetitive elements during seed development underlies gene imprinting.** *Science* 2009, **324:**1447-1451.

5. Mosher RA, Melnyk CW, Kelly KA, Dunn RM, Studholme DJ, Baulcombe DC: **Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*.** *Nature* 2009, **460:**283-286.

6. Hsieh TF, Ibarra CA, Silva P, Zemach A, Eshed-Williams L, Fischer RL, Zilberman D: **Genome-wide demethylation of *Arabidopsis* endosperm.** *Science* 2009, **324:**1451-1454.

7. Yadegari R, Drews GN: **Female gametophyte development.** *Plant Cell* 2004, **16 Suppl:**S133-S141.

8. McCormick S: **Control of male gametophyte development.** *Plant Cell* 2004, **16 Suppl:**S142-S153.

9. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL: **DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*.** *Cell* 2002, **110:**33-42.

10. Huh JH, Bauer MJ, Hsieh TF, Fischer RL: **Cellular programming of plant gene imprinting.** *Cell* 2008, **132:**735-744.

11. Makarevich G, Villar CB, Erilova A, Kohler C: **Mechanism of *PHERES1* imprinting in *Arabidopsis*.** *J Cell Sci* 2008, **121:**906-912.

12. Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J: **Transgenerational stability of the *Arabidopsis* epigenome is coordinated by CG methylation.** *Cell* 2007, **130:**851-862.

13. 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:**6853-6858.

14. AGI: **Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*.** *Nature* 2000, **408:**796-815.

15. Oakeley EJ, Podesta A, Jost JP: **Developmental changes in DNA methylation of the two tobacco pollen nuclei during maturation.** *Proc Natl Acad Sci USA* 1997, **94:**11721-11725.

16. Lauria M, Rupe M, Guo M, Kranz E, Pirona R, Viotti A, Lund G: **Extensive maternal DNA hypomethylation in the endosperm of *Zea mays*.** *Plant Cell* 2004, **16:**510-522.

17. Haberer G, Young S, Bharti AK, Gundlach H, Raymond C, Fuks G, Butler E, Wing RA, Rounsley S, Birren B, Nusbaum C, Mayer KF, Messing J: **Structure and architecture of the maize genome.** *Plant Physiol* 2005, **139:**1612-1624.

18. Lisch D: **Epigenetic regulation of transposable elements in plants.** *Annu Rev Plant Biol* 2009, **60:**43-66.