24. B. Brower-Toland et al., Genes Dev. 21, 2300 (2007).
25. M. Pal-Bhadra et al., Science 303, 669 (2004).
26. C. Klattenhoff et al., Cell 138, 1137 (2009).
27. S. D. Taverna, R. S. Coyne, C. D. Allis, Cell 110, 701 (2002).
28. A. Böhne, F. Brunet, D. Galiana-Arnoux, C. Schultheis, J. N. Volff, Chromosome Res. 16, 203 (2008).
29. S. Chambeyron et al., Proc. Natl. Acad. Sci. U.S.A. 105, 14964 (2008).
30. C. Josefsson, B. Dilkes, L. Comai, Curr. Biol. 16, 1322 (2006).
31. R. A. Martienssen, New Phytol. 186, 46 (2010).
32. A. A. Aravin, R. Sachidanandam, A. Girard, K. Fejes-Toth, G. J. Hannon, Science 316, 744 (2007).

33. N. Zamudio, D. Bourc'his, Heredity 105, 92 (2010).
34. H. Kano et al., Genes Dev. 23, 1303 (2009).
35. R. J. O'Neill, M. J. O'Neill, J. A. Graves, Nature 393, 68 (1998).
36. C. Babinet, V. Richoux, J. L. Guénet, J. P. Renard, Dev. Suppl. 1990, 81 (1990).
37. T. A. Bell et al., Genetics 172, 411 (2006).
38. T. Watanabe et al., Nature 453, 539 (2008).
39. W. Yan et al., Biol. Reprod. 78, 896 (2008).
40. V. Grandjean et al., Development 136, 3647 (2009).
41. Y. Tarutani et al., Nature 466, 983 (2010).
42. L. Wu et al., Mol. Cell 38, 465 (2010).

43. P. Dunoyer et al., EMBO J. 29, 1699 (2010).
44. X. Dramard, T. Heidmann, S. Jensen, PLoS ONE 2, e304 (2007).
45. We are indebted to J. Brennecke, S. Duharcourt, and A. Pélisson for valuable discussions and advice. D.B. is supported by a European Young Investigator award and the Fondation Schlumberger pour l'Enseignement et la Recherche. O.V. is supported by a European Research Council starting grant (210890) "Frontiers of RNAi" and an award from the Bettencourt Foundation.

REVIEW

Epigenetic Reprogramming in Plant and Animal Development

Suhua Feng,[1] Steven E. Jacobsen,[1]* Wolf Reik[2]*

Epigenetic modifications of the genome are generally stable in somatic cells of multicellular organisms. In germ cells and early embryos, however, epigenetic reprogramming occurs on a genome-wide scale, which includes demethylation of DNA and remodeling of histones and their modifications. The mechanisms of genome-wide erasure of DNA methylation, which involve modifications to 5-methylcytosine and DNA repair, are being unraveled. Epigenetic reprogramming has important roles in imprinting, the natural as well as experimental acquisition of totipotency and pluripotency, control of transposons, and epigenetic inheritance across generations. Small RNAs and the inheritance of histone marks may also contribute to epigenetic inheritance and reprogramming. Reprogramming occurs in flowering plants and in mammals, and the similarities and differences illuminate developmental and reproductive strategies.

Epigenetic marks are enzyme-mediated chemical modifications of DNA and of its associated chromatin proteins. Although they do not alter the primary sequence of DNA, they also contain heritable information and play key roles in regulating genome function. Such modifications—including cytosine methylation, posttranslational modifications of histone tails and the histone core, and the positioning of nucleosomes (histone octamers wrapped with DNA)—influence the transcriptional state and other functional aspects of chromatin. For example, methylation of DNA and certain residues on the histone H3 N-terminal tail [e.g., H3 lysine 9 (H3K9)] are important for transcriptional gene silencing and the formation of heterochromatin. Such marks are essential for the silencing of nongenic sequences—including transposons, pseudogenes, repetitive sequences, and integrated viruses—that could become deleterious to cells if expressed and hence activated. Epigenetic gene silencing is also important in developmental phenomena such as imprinting in both plants and mammals, as well as in cell differentiation and reprogramming.

DNA methylation occurs in three different nucleotide sequence contexts: CG, CHG, and CHH (where H = C, T, or A). In both mammals and plants, CG methylation is maintained by the maintenance DNA methyltransferase, termed DNMT1 [DNA (cytosine-5)-methyltransferase 1] in mammals and MET1 (DNA METHYLTRANSFERASE 1) in Arabidopsis, and by a cofactor that recognizes hemimethylated DNA at replication foci, called UHRF1 (ubiquitin-like containing PHD and RING finger domains 1) in mammals and VIM (VARIATION IN METHYLATION) family proteins in Arabidopsis (1). In addition, the mammalian de novo DNA methyltransferases DNMT3A and Dnmt3b are required for the maintenance of CG methylation at some loci (2). CHG methylation is common in Arabidopsis and other plant genomes and is maintained by a feedforward loop that is formed by a plant-specific DNA methyltransferase, CMT3 (CHROMOMETHYLASE 3), and a histone methyltransferase, KYP (KRYPTONITE) (1, 3, 4). CHH methylation is also abundant in plants and is maintained by the RNA-directed DNA methylation (RdDM) pathway, which actively targets the DNA methyltransferase DRM2 (DOMAINS REARRANGED METHYLTRANSFERASE 2; a homolog of Dnmt3) to DNA by means of 24-nucleotide (nt) small interfering RNAs (siRNAs) bound by AGO4 (ARGONAUTE 4) (1) (Fig. 1). CHG and CHH methylation are also present at detectable levels in mammals, especially in stem cells, and this methylation is likely introduced by DNMT3A and DNMT3B (5, 6). De novo methylation of DNA in all of these sequence contexts is generally established by the DNMT3 (mammals) and DRM2 (Arabidopsis) methyltransferases. Mammals do not have an Arabidopsis-like RNA-directed DNA methylation pathway, but in germ cells, PIWI-associated RNAs (piRNAs) are thought to guide DNMT3 activity (7). Mammals have a noncatalytic paralog of de novo methyltransferase, DNMT3L, which interacts with DNMT3A and unmethylated H3K4 (as does DNMT3A and DNMT3B) (8–10); these findings imply a targeting mechanism of these methyltransferases to chromatin. Unmethylated CpG islands are specifically bound by CFP1 (CXXC finger protein–1), which in turn recruits the histone H3K4 methyltransferase SETD1 (SET domain containing–1) (11); this suggests that H3K4 methylation, and therefore exclusion of DNMT3 from CpG islands, could help to explain how promoters remain unmethylated. Consistent with this idea, demethylation of H3K4 has been shown to be important for the acquisition of DNA methylation in imprinted genes in oocytes (12). Additionally, transcription can also help to establish de novo DNA methylation at imprinted regions (13). Earlier this year, it was shown that the nucleosome landscape also influences the methylation patterning in both plant and animal genomes (14).

Some histone modifications are also thought to be actively maintained during DNA replication, in part facilitated by the association of the histone modification enzymes with the DNA replication machinery. For example, the mammalian histone H3K9 methyltransferases G9A and SETDB1 (SET domain bifurcated–1), the mammalian H4K20 methyltransferase SETD8 (SET domain containing–8), and the plant histone H3K27 monomethyltransferases ATXR5 (ARABIDOPSIS TRITHORAX-RELATED PROTEIN 5) and ATXR6 interact with the replication protein PCNA (proliferating cell nuclear antigen) (15, 16). However, histone methylation can also be very dynamic and is controlled

[1]Howard Hughes Medical Institute and Department of Molecular, Cell and Developmental Biology, University of California, Los Angeles, CA 90095, USA. [2]Laboratory of Developmental Genetics and Imprinting, Babraham Institute, Cambridge CB22 3AT, UK, and Centre for Trophoblast Research, University of Cambridge, Cambridge CB2 3EG, UK.

*To whom correspondence should be addressed. E-mail: jacobsen@ucla.edu (S.E.J.); wolf.reik@bbsrc.ac.uk (W.R.)

by the combined action of both histone methyltransferases and histone demethylases, as well as by the proteins that read these histone marks (16).

Through the developmental regulation of these epigenetic mechanisms, both plants and animals undergo epigenetic reprogramming in various cell types and developmental stages, which serve either to transmit epigenetic information between cells or between sexual generations, or to reset epigenetic marks to reduce the risk of perpetuating dangerous epigenetic alleles.

DNA Methylation Throughout the *Arabidopsis* Life Cycle

To maintain genome integrity from generation to generation, transposons and repetitive DNA elements must be kept under tight regulation in reproductive cells. One of the ways that plants achieve this is through the stable inheritance of DNA methylation. Plants frequently show meiotic inheritance of gene silencing (1). Furthermore, plants are not known to undergo genome-wide waves of demethylation in germ cells, as occurs in animals. However, large-scale reprogramming occurs in non–germ line reproductive cells, and this reprogramming may function to reinforce silencing of transposable elements in germ cells (see below).

One way to actively reprogram the epigenome is to remove methylated cytosines. The *Arabidopsis* genome encodes four bifunctional helix-hairpin-helix DNA glycosylases and AP lyases—ROS1 (REPRESSOR OF SILENCING 1), DME (DEMETER), DML2 (DEMETER-LIKE 2), and DML3 (DEMETER-LIKE 3)—which recognize and remove methylated cytosines, resulting in a 1-nt gap in the DNA double strand. Subsequently, as yet unidentified DNA repair polymerase and DNA ligase enzymes are thought to fill in the gap with an unmethylated cytosine (1, 17). ROS1, DML2, and DML3 mainly function in vegetative tissues, and genomic studies suggest that they demethylate hundreds of specific loci across the genome with a bias toward genes (1, 18). Knocking out all three genes does not markedly affect the overall levels or patterns of methylation in the *Arabidopsis* genome (18, 19). Instead, these enzymes appear to be acting as a counterbalance to the RNA-directed DNA methylation system to quantitatively fine-tune methylation levels at particular genomic locations.

By contrast, DME functions to cause global hypomethylation in the endosperm (the extra-embryonic tissue of flowering plants) of *Arabidopsis* (20, 21), and thus contributes to large-scale epigenetic reprogramming (Fig. 1). In *Arabidopsis*, female gametogenesis begins when a somatically derived megaspore mother cell undergoes meiosis to give rise to a haploid megaspore, which subsequently develops into a mature female gametophyte (embryo sac) that contains one egg cell, one central cell (two nuclei), and several other accessory cells. During double fertilization (which is common in plants), the egg cell fuses with a sperm cell from the male gametophyte (pollen grain) to form an embryo, and the central cell fuses with the other sperm cell from pollen to form the triploid endosperm, which nourishes the embryo, and thus bears a function similar to that of the placenta in mammals. DME is expressed primarily in the central cell before fertilization, and thus only the maternal genome is demethylated by DME. This leads to maternal allele–specific gene expression (imprinting) in the endosperm (22). Until recently, only six imprinted *Arabidopsis* genes were known, but recent genomic studies of endosperm have revealed genome-wide differences in DNA methylation, including a substantial reduction of CG methylation; hence, many additional genes are likely to be imprinted in *Arabidopsis*, some of which have been verified by single-gene studies (20, 21) (Fig. 1). Demethylation by DME may also reactivate transposon expression, which shunts transposon transcripts into the RNAi pathway, producing additional siRNAs that can guide DNA methylation to non-CG sites whose methylation is high in wild-type endosperm but decreased in *dme* mutant endosperm (Fig. 1). Curiously, there are even higher levels of non-CG methylation in the wild-type embryo that could be explained by movement of siRNAs produced in the central cell into the egg cell; this attractive idea awaits experimental support (20). Because the endosperm genome does not contribute to the next generation, mild reactivation of transposons in endosperm may not be deleterious and may reinforce the silencing of transposons in the egg cell and later embryo, contributing to the genome integrity of offspring. Indeed, there is a class of RNA polymerase IV (Pol IV)–dependent siRNAs that only accumulates in flowers and young siliques, likely originating from the endosperm (23). Notably, these siRNAs are derived from maternal alleles only, which suggests that they may be produced in part during female gametogenesis and then retained after karyogamy. However, these siRNAs are expressed more highly after fertilization, and therefore imprinted maternal expression of siRNA loci may also occur as the endosperm develops (23). It is tempting to speculate that the maternal Pol IV–dependent siRNAs are the "messenger" that mediates communication between endosperm and embryo (Fig. 1); however, these siRNAs were detected only in the endosperm, not in the embryo (23). Nonetheless, the possibility that they exist in low abundance in the embryo, or are ephemeral, cannot be ruled out.

The idea that siRNAs move from the endosperm to the embryo is consistent with the model put forth in an earlier study on paternal genome reprogramming in *Arabidopsis* (24). The male gametophyte of *Arabidopsis* (a pollen grain) contains two sperm cells, which fertilize the egg cell and central cell, respectively, and a vegetative nucleus (Fig. 1). Transposon expression is generally up-regulated in pollen, and certain transposons even become mobile in pollen, unlike the situation in most other tissues (24). Reduction of transposon methylation followed by transposon reactivation appears to occur in the vegetative nucleus; this is supported by the finding that transposon reactivation and movement are not inherited by the next generation (24). It has been shown that several key RdDM pathway proteins (RDR2 and DCL3) and CHG methylation maintenance pathway proteins (CMT3 and KYP) have reduced expression levels in pollen; in addition, DDM1 (DECREASE IN DNA METHYLATION 1), an important chromatin remodeler required for DNA and histone methylation and transposon silencing, is exclusively localized in sperm cells but not in the vegetative nucleus (24). These results suggest a model in which hypomethylation of the vegetative cell may reactivate transposons that could serve to reinforce transposon silencing in the adjacent sperm cells (Fig. 1) (24).

Small RNAs may be involved in this communication between the vegetative cell and the sperm cells. A class of siRNAs that is 21 nt in length and corresponds to *Athila* retrotransposons, the largest transposon family, is detected in sperm cells. Because *Athila* retrotransposons remain silenced in sperm cells but are activated in the vegetative nucleus, it is possible that the 21-nt siRNAs are produced in the vegetative nucleus and then travel to their site of action—sperm cells—where they mediate the silencing of transposons through an unknown mechanism (Fig. 1) (24). A common theme is that both male and female gametophytes contain nurse cells in which massive epigenetic reprogramming may serve to reinforce transposon silencing in the germ line (Fig. 1).

Another example of small RNAs silencing transposons at a distance occurs when the megaspore mother cell differentiates from somatic tissues (25). Mutations in AGO9 (ARGONAUTE 9), a member of the *Arabidopsis* Argonaute family of proteins, result in the reactivation of transposons in the ovule (including the egg cell) (Fig. 1). Remarkably, *AGO9* is not expressed in the reproductive cells themselves (megaspore mother cell, megaspore, or developing female gametophyte), but is expressed in the companion cells surrounding the female gametophyte. The transposon targets of AGO9 are similar to those reactivated in pollen, and evidence suggests that AGO9-mediated transposon silencing uses components of known silencing pathways, including the 24-nt RNA-directed DNA methylation pathway (25). Whether the AGO9-associated 24-nt siRNAs are the mobile signal remains to be tested.

Resetting of Histone Modifications in *Arabidopsis*

In addition to DNA methylation, plants also reprogram histones and their associated marks;
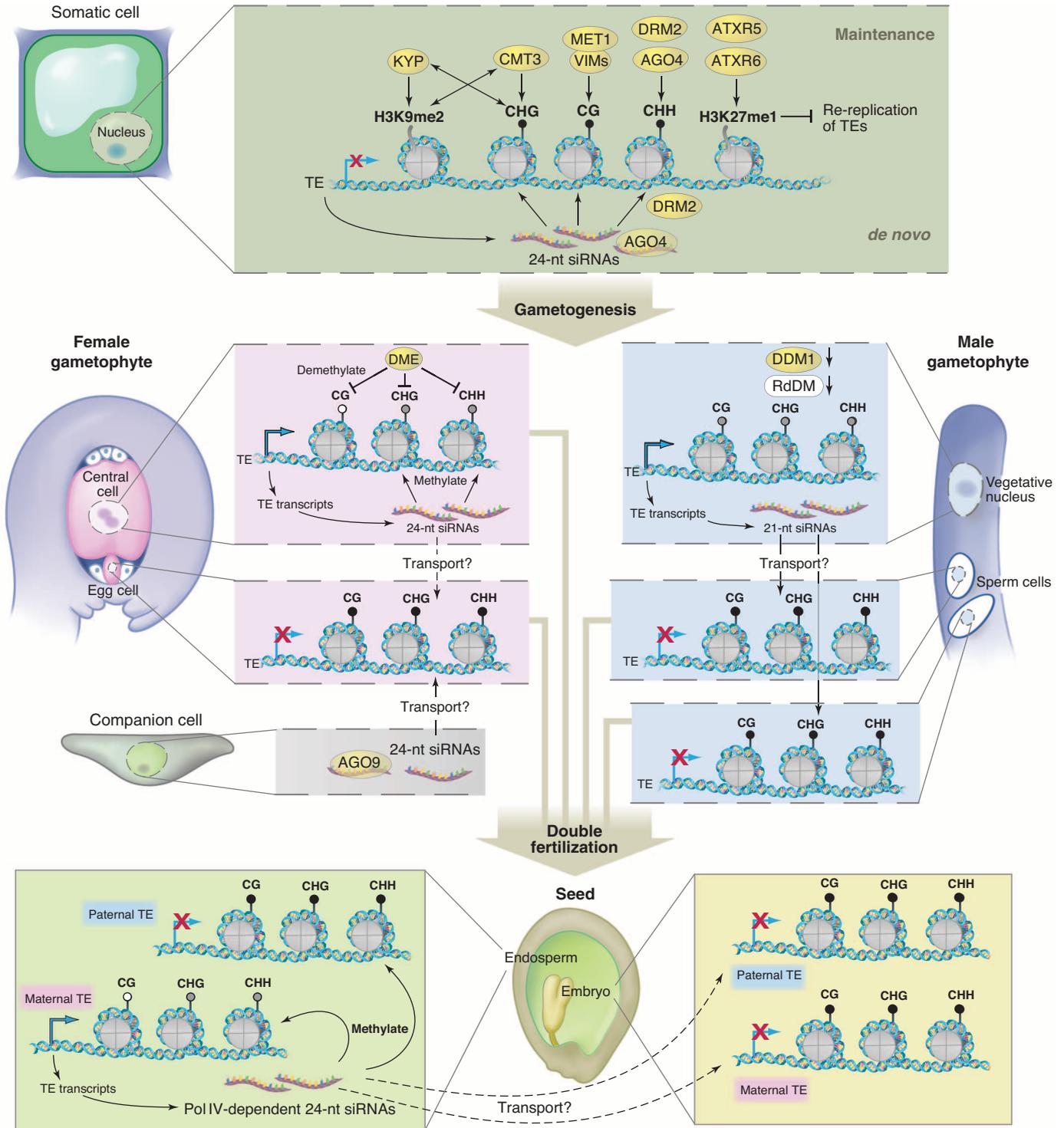
Epigenetics

Fig. 1. Model of epigenetic silencing dynamics during the *Arabidopsis* life cycle. In somatic cells, three different mechanisms are responsible for repressing transcription from transposable elements (TEs), DNA methylation (in all three sequence contexts), histone H3K9 dimethylation (H3K9me2), and histone H3K27 monomethylation (H3K27me1). Methyltransferases and proteins regulating these epigenetic marks are shown. See text for details. In the female gametophyte, the central cell is demethylated by DME, which leads to TE activation and up-regulation of RdDM. The siRNAs produced from TEs not only direct non-CG methylation in the central cell, but also might travel to the egg cell and enhance the silencing of TEs there. In addition, AGO9-associated siRNAs produced in somatic companion cells also contribute to the silencing of TEs in the egg cell. In the male gametophyte, the vegetative nucleus does not express DDM1 and has reduced RdDM, which leads to TE activation and mobilization. A new class of 21-nt siRNAs, produced from TEs in the vegetative nucleus, travels to sperm cells to reinforce TE silencing. After double fertilization, maternal TEs in the endosperm stay activated and produce Pol IV–dependent siRNAs, which could function to silence the paternal TEs in the endosperm. The methylation levels in the embryo are elevated, possibly as a result of the siRNA signals transmitted from the endosperm. Different shadings indicate the level of DNA methylation (high, black; medium, gray; low, white).

as opposed to DNA methylation (which is typically inherited), some histone modifications are known to be reset in each generation. Because plants do not set aside a germ line early in development (germ cells are differentiated from adult somatic cells), some type of "reprogramming" process is likely needed to erase the effects of epigenetic marks caused by external stimuli (such as development or stress). For example, PCG (Polycomb group) proteins mediate the silencing of *FLC* (*FLOWERING LOCUS C*) in *Arabidopsis*, which controls flowering time (*26*). In winter-annual accessions of *Arabidopsis*, *FLC* is expressed at high levels to repress the initiation of flowering. During vernalization (prolonged exposure to cold, such as during winter), *FLC* becomes modified by H3K27 trimethylation, which helps to turn off *FLC* expression epigenetically. When winter passes and temperatures become warmer, trimethylation and silencing of *FLC* persists, and therefore *Arabidopsis* can flower in response to environmental cues such as photoperiod. When gametes are formed through meiosis, H3K27 trimethylation marks on *FLC* are removed by an unknown mechanism and *FLC* becomes reexpressed in the seeds. Thus, flowering is inhibited by *FLC* until the next-generation plants encounter cold weather.

Resetting of histone marks may involve, in part, global replacement of histones (*27*). The histone variant H3.3 can be incorporated in the absence of DNA replication, and thus is a candidate for the "replacement" histone H3 during reprogramming. HTR10 (HISTONE THREE RELATED 10) is exclusively expressed in male reproductive cells, but after karyogamy of sperm and egg cell nuclei, the paternal HTR10 signal disappears within a matter of hours before S phase of the first zygote division (*27*). This suggests that HTR10 is actively removed from the chromatin in a replication-independent manner specifically in the sperm cell that fertilizes the egg. Unlike DNA methylation reprogramming, which occurs in accessory cells, histone reprogramming takes place in the zygote and thus can transmit information to the next generation. These results raise a number of questions. How does the reprogramming system differentiate between the two sperm cells? Does similar reprogramming happen in the female genome as well? What types of histone H3 replace the parental histone H3 in the zygote, and where do they come from?

Recently a new transposon silencing mark was described in *Arabidopsis* that does not appear to involve the well-studied DNA methylation or histone H3K9 dimethylation marks. This mark, H3K27 monomethylation, is needed to suppress excessive replication of heterochromatin in which transposons reside (*15*). Overreplication of transposons might lead to transposon reactivation and copy number propagation, and the H3K27 monomethylation system may have

evolved to suppress excessive replication and to ensure genome stability (Fig. 1). If true, this would suggest that histone marks not only get reprogrammed but also reprogram the genome, in the case of H3K27 monomethylation, by keeping the replication of transposons in check. This presumably is important for actively cycling plant cells, for reproductive cells undergoing meiosis, and perhaps for early stages of embryo development.

Mechanisms of Epigenetic Reprogramming in Mammalian Development

Genome-wide epigenetic reprogramming occurs in mammalian development at two distinct stages: in primordial germ cells (PGCs) primarily once they have reached the embryonic gonads (embryonic day E10.5 to E13.5), and in the early embryo beginning in the zygote immediately after fertilization and extending to the morula stage of preimplantation development (Fig. 2) (*28–30*). This reprogramming entails erasure of DNA methylation and loss of histone modifications (as well as loss of histones and histone variants); here we focus on demethylation of DNA. The loss of DNA methylation by E13.5 (the developmental endpoint of reprogramming) is truly global; in mouse female PGCs, only 7% of CpGs remain methylated [versus 70 to 80% in embryonic stem (ES) cells and somatic cells], and most promoters and genic, intergenic, and transposon sequences are hypomethylated at this stage (*31*). The only clear exception to global erasure is intracisternal A particles (IAPs), an active family of retrotransposons that have only recently been acquired in the rodent lineage, which only show partial demethylation in PGCs (*31*). Promoters of germ cell–specific genes (such as *Dazl* or *Vasa*) are methylated in early PGCs and become demethylated and expressed during reprogramming (*32*). Imprinted genes have allele-specific methylation in early PGCs and the *Xist* promoter is methylated, and this methylation is all erased in PGCs by E13.5 (Fig. 2) (*33, 34*). Although most of the genome-wide demethylation appears to occur in E11.5 to E13.5 PGCs, it remains possible that some loci become demethylated at slightly earlier stages (*35*); hence, demethylation is not necessarily coordinated timewise throughout the genome. Nothing is currently known about the possible occurrence or erasure of non-CG methylation in PGCs.

DNA deaminases and the base excision repair pathway have recently been implicated in erasure, which suggests that active demethylation is involved at least in part (*31, 36*). The cytosine deaminases AID and APOBEC1 are capable in vitro of deaminating 5-methylcytosine (5mC) as well as cytosine and are expressed, albeit at a low level, in PGCs (*36, 37*). Notably, AID deficiency in PGCs results in a deficit in demethylation of 20% of all CpGs if it is assumed that early PGCs have methylation levels similar to those of ES

cells or somatic cells (*31*). Because of this partial effect of AID deficiency on erasure, the potential redundancy with other DNA deaminases needs to be examined. The 5mC hydroxylases TET1 and TET2 are also expressed in PGCs (*36*), suggesting the possibility that 5mC could be modified by different mechanisms (deamination, hydroxylation) in order to initiate active demethylation. It is also possible that a combination of passive (resulting in hemimethylated substrates in G_2 phase of the cell cycle) and active demethylation could be involved. Finally, it is possible that the genome-wide nature of the demethylation process and its relatively coordinate timing require different mechanisms and different modifications of 5mC to join forces in order to achieve such large-scale reprogramming.

Initial modification of 5mC would require further modification or DNA repair in order to achieve demethylation. DNA repair pathways that might be involved in resolving mismatches or in excising 5-hydroxymethylcytosine (5hmC) are nucleotide excision repair, mismatch repair, and especially base excision repair (BER), which is also involved in demethylation during reprogramming in plants (*1*). BER components such as PARP1, APE1, and XRCC1 are all up-regulated at E11.5 in PGCs, together with enhancement of chromatin-bound XRCC1; thus, it is possible that BER is activated at this time point (Fig. 2) (*36*). Global losses of several histone modifications (e.g., H3K27me3, H3K9ac) as well as the linker histone H1 are observed after demethylation of DNA, indicating that widespread DNA repair might be associated with global remodeling of nucleosomes in PGCs (*38*). It is also possible that specific histone modification or demodification enzymes (deacetylases, demethylases) are in part responsible for erasure of histone marks in PGCs, but none have been identified so far.

Base excision repair also appears to be involved in demethylation in the zygote immediately after fertilization (Fig. 2). The added complication here is that it is specifically the paternal, sperm-derived, genome that is demethylated, whereas the maternal one is not; the maternal genome may be specifically protected from demethylation (*39–42*). Differentially methylated regions in imprinted genes are also specifically protected from demethylation, and so again are IAPs. Nonetheless, there appear to be substantial losses of methylation in the zygote, potentially of a similar scale to those occurring in PGCs (*39, 43, 44*). Notably, demethylation of the paternal genome may occur in two phases, one before DNA replication and one associated with the S and G_2 phases (*44*). The first phase might involve modification of 5mC but only partial demethylation (*44*). Demethylation might then continue at replication or afterward. BER components are also present at these stages with an

enhancement of chromatin-bound XRCC1 in the paternal pronucleus (*36*). Both phases show evidence of DNA strand breaks, indicating that repair may be involved in both of them, and inhibition of BER components partly interferes with demethylation (*36, 44*). Whether AID or TETs are involved in zygotic demethylation is not yet known, but components of the Elongator complex (Elongator complex proteins, ELPs)

the blastocyst stage, with DNMT1 protein being largely excluded from the nucleus by an unknown mechanism (Fig. 2). Nonetheless, the maintenance of methylation in differentially methylated regions of imprinted genes does depend on DNMT1 (*46*), so it will be important to understand how DNMT1 might be targeted during this reprogramming phase to key regions in the genome, such as imprinted genes (*47*).

versely, homologs of APOBEC deaminases and TET and ALKBH-type hydroxylases have yet to be described in plants.

Experimental Reprogramming in Mammals

Experimental reprogramming to a pluripotent state can be achieved, albeit inefficiently, by fusion of somatic cells and pluripotent cells, by cloning, and by direct reprogramming using core transcription factors (*48*). With all three methods, there is evidence that epigenetic reprogramming is a central component of achieving the goal of an embryonic or ES cell–like (iPS) state. In cell fusion experiments between somatic cells and ES cells, key pluripotency genes such as *Oct4* and *Nanog* need to be demethylated; AID also seems to be important for demethylation in this system (*49*). Generation of iPS cells from somatic cells by the transduction of core transcription factors (such as OCT4, SOX2, KLF4, and C-MYC) probably requires multiple epigenetic reprogramming steps while the cells that undergo reprogramming divide (*50*). DNA demethylation is clearly critical because incompletely reprogrammed iPS cells can become completely reprogrammed by treatment with the methylation inhibitor azacytidine (*48*). Inhibitors of histone deacetylases and histone methyltransferases are also beneficial, showing in general that repressive epigenetic modifications acquired during differentiation and somatic development need to be reversed to achieve the pluripotent state (*48*). Notably, reprogramming by cloning apparently results in better resetting of the epigenome than can be achieved by direct reprogramming with transcription factors, indicating perhaps that true totipotency requires passage through germ cells or zygotes (*51*). Direct applications to regenerative medicine will result from unraveling the role of AID, hydroxymethylation, and the TETs, and of the base excision repair pathway as well as the methyltransferases in this process, and from knowledge of how the reprogramming network is connected with the pluripotency network.

Fig. 2. The two major phases of genome-wide erasure of DNA methylation in the early embryo and in primordial germ cells (PGCs) of the mouse. Thickness of the outer arrows indicates levels of DNA methylation. Red, maternal genome; blue, paternal genome. After fertilization, the paternal genome is more rapidly demethylated than the maternal one. During gametogenesis, de novo methylation in spermatogenesis occurs earlier than in oogenesis. The inner circle shows factors or candidate factors that are implicated in de novo methylation, the maintenance of methylation, and demethylation, respectively. Solid arrows in the inner circle show at what developmental time these epigenetic regulators are thought to act. ES cells, TS cells, and XEN cells are stem cell lines that are derived from the inner cell mass (ICM), trophectoderm (TE), and primitive endoderm (PE) of the blastocyst, respectively.

have been implicated in demethylation of the paternal genome (Fig. 2) (*45*); Elongator is involved in diverse aspects of transcriptional regulation and can also modify tRNAs. Could Elongator catalyze an as yet unknown modification of 5mC that makes it a substrate for BER? After zygotic demethylation, the embryonic genome continues to be demethylated during the following few cleavage divisions until

Hence, the current evidence for the initiation and regulation of genome-wide erasure of DNA methylation in PGCs and the zygote points to initiating events that modify 5mC (such as deamination and hydroxylation), which would trigger a BER response. Of course, it is still possible that bifunctional DNA glycosylases of the type that excise 5mC in plants also exist in animals (although none have been found so far); con-

Comparative Biology of Epigenetic Reprogramming

Whether genome-scale epigenetic reprogramming has a unified purpose is not clear; some aspects of reprogramming are clearly conserved (or have been reinvented) in animals and plants with their contrasting, although sometimes surprisingly similar, reproductive and biological strategies. In mammals, zygotic reprogramming is broadly conserved, although there may be some differences in timing or extent; by contrast, *Xenopus* does not appear to show demethylation of the paternal genome (*52*). Hypomethylation of PGCs is also seen in human and pig fetal development but has not been studied in nonmammalian organisms. Global DNA demethylation in PGCs and paternal demethylation in the
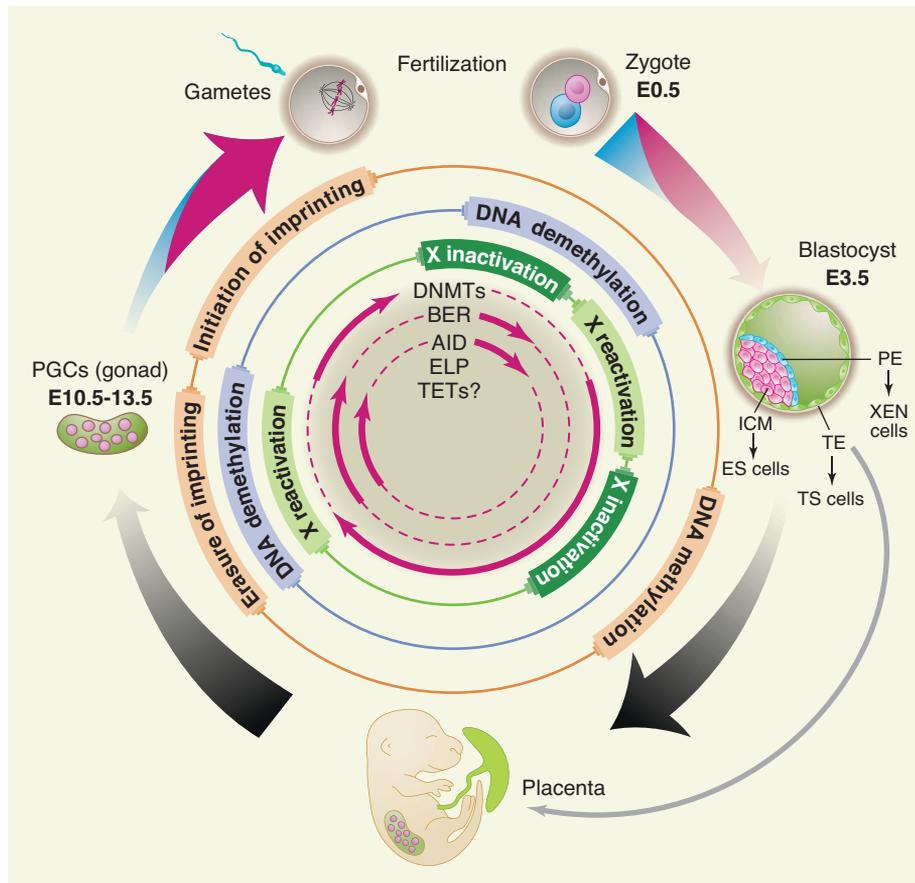
zygote may occur primarily in mammals (and in the central cell in seed plants) that also have imprinting whose mechanism is based on DNA methylation. Clearly, demethylation in PGCs is necessary for erasure of imprints so that new imprints can later be established properly, according to the sex of the germ line (Fig. 2). Plants do not seem to erase imprints; instead, they establish them by demethylation of the maternal genome in the endosperm after fertilization (with the endosperm being comparable to the placenta) (Fig. 1). Perhaps there are as yet undiscovered imprinted genes that acquire parent-specific methylation patterns by (paternal) zygotic demethylation, in analogy to plants.

A second group of genes where demethylation in PGCs seems important are the germ line–specific genes (e.g., *Dazl, Vasa*) that have specialized functions, for example, in meiosis and germ cell differentiation. These genes are generally demethylated and expressed in germ cells, but in early PGCs they are methylated and silenced. Genes that are demethylated in PGCs include those with a role in transposon control; *Tex19.1*, for example, silences members of the ERVK transposon family (*53*). Hence, global demethylation, which in principle would lead to transcriptional activation and potentially to transposition of active transposon families, at the same time activates defense mechanisms against transposons that are not needed in somatic cells where transposons are methylated. An extreme view of this scenario is the possibility that demethylated transposons produce small RNAs, which in turn lead to de novo methylation and renewed silencing of transposons (Fig. 1) (*24*). Although it may sound paradoxical, reprogramming may have an important role in resetting the permanent silencing program for transposons across generations. Also, the fact that AID has a role in the erasure of methylation in PGCs is interesting in connection with roles of APOBEC deaminases in innate immunity and transposon control, establishing another potential link between the two.

The extent of methylation reprogramming in PGCs is substantial, and this limits the potential in mammals for epigenetic transgenerational inheritance. By contrast, in plants where epigenetic reprogramming may not occur to such an extent in the germ line, examples of stable inheritance of epialleles over multiple generations are more common (*54*). In *Caenorhabditis elegans*, histone demethylation in the germ line is needed to prevent accumulation of aberrant epigenetic marks that interfere with normal physiology and limit life span (*55*). By analogy, epigenetic reprogramming in mammalian PGCs or early embryos may be important to prevent the accumulation of potentially detrimental epialleles, which could otherwise cause chronic diseases and limit life span in human populations. The inheritance of histone marks and of small

RNAs potentially through both oocyte and sperm might also contribute to epigenetic inheritance and to reprogramming across generations in animals and in plants (*24, 25, 56–59*).

Finally, epigenetic reprogramming is linked to regaining pluripotency and, following that, lineage commitment. Early PGCs and cells in the early embryo are pluripotent, and these cells as well as the stem cell lines that can be isolated from them [ES and EG (embryonic germ) cells] have unique epigenetic signatures, which are at least in part the outcomes of reprogramming. For example, some of the key pluripotency transcription factors (such as *Nanog*) are methylated in sperm but not in ICM (inner cell mass) or ES cells, so their demethylation is important for the acquisition of pluripotency (*43*); in the absence of the highly expressed gene encoding TET1 (which hydroxylates 5mC), the *Nanog* gene is repressed and its promoter becomes methylated (*60*).

One final epigenetic parallel between mammals and plants is worth highlighting. After demethylation in the early mammalian embryo, selective de novo methylation occurs in ICM cells and their descendants, which is important for the identity and stability of embryonic lineages (*28*), whereas the placenta remains hypomethylated at the genome-wide level (*31*). Similarly, genome-wide demethylation in the plant endosperm but not the embryo (*20, 21*) indicates that epigenetic regulation between the two primary lineages (embryonic, extraembryonic) is fundamentally different, with this difference apparently being conserved—or reinvented—in plants and animals.

References and Notes

1. J. A. Law, S. E. Jacobsen, *Nat. Rev. Genet.* **11**, 204 (2010).
2. T. Chen, Y. Ueda, J. E. Dodge, Z. Wang, E. Li, *Mol. Cell. Biol.* **23**, 5594 (2003).
3. S. Feng *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8689 (2010).
4. A. Zemach, I. E. McDaniel, P. Silva, D. Zilberman, *Science* **328**, 916 (2010); published online 15 April 2010 (10.1126/science.1186366).
5. B. H. Ramsahoye *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5237 (2000).
6. R. Lister *et al.*, *Nature* **462**, 315 (2009).
7. A. A. Aravin, G. J. Hannon, *Cold Spring Harb. Symp. Quant. Biol.* **73**, 283 (2008).
8. D. Jia, R. Z. Jurkowska, X. Zhang, A. Jeltsch, X. Cheng, *Nature* **449**, 248 (2007).
9. S. K. Ooi *et al.*, *Nature* **448**, 714 (2007).
10. Y. Zhang *et al.*, *Nucleic Acids Res.* **38**, 4246 (2010).
11. J. P. Thomson *et al.*, *Nature* **464**, 1082 (2010).
12. D. N. Ciccone *et al.*, *Nature* **461**, 415 (2009).
13. M. Chotalia *et al.*, *Genes Dev.* **23**, 105 (2009).
14. R. K. Chodavarapu *et al.*, *Nature* **466**, 388 (2010).
15. Y. Jacob *et al.*, *Nature* **466**, 987 (2010).
16. R. Bonasio, S. Tu, D. Reinberg, *Science* **330**, 612 (2010).
17. J. K. Zhu, *Annu. Rev. Genet.* **43**, 143 (2009).
18. J. Penterman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6752 (2007).
19. R. Lister *et al.*, *Cell* **133**, 523 (2008).
20. T. F. Hsieh *et al.*, *Science* **324**, 1451 (2009).
21. M. Gehring, K. L. Bubb, S. Henikoff, *Science* **324**, 1447 (2009).
22. J. H. Huh, M. J. Bauer, T. F. Hsieh, R. L. Fischer, *Cell* **132**, 735 (2008).
23. R. A. Mosher *et al.*, *Nature* **460**, 283 (2009).
24. R. K. Slotkin *et al.*, *Cell* **136**, 461 (2009).
25. V. Olmedo-Monfil *et al.*, *Nature* **464**, 628 (2010).
26. D. H. Kim, M. R. Doyle, S. Sung, R. M. Amasino, *Annu. Rev. Cell Dev. Biol.* **25**, 277 (2009).
27. M. Ingouff, Y. Hamamura, M. Gourgues, T. Higashiyama, F. Berger, *Curr. Biol.* **17**, 1032 (2007).
28. M. Hemberger, W. Dean, W. Reik, *Nat. Rev. Mol. Cell Biol.* **10**, 526 (2009).
29. M. A. Surani, K. Hayashi, P. Hajkova, *Cell* **128**, 747 (2007).
30. H. Sasaki, Y. Matsui, *Nat. Rev. Genet.* **9**, 129 (2008).
31. C. Popp *et al.*, *Nature* **463**, 1101 (2010).
32. D. M. Maatouk *et al.*, *Development* **133**, 3411 (2006).
33. P. Hajkova *et al.*, *Mech. Dev.* **117**, 15 (2002).
34. J. Lee *et al.*, *Development* **129**, 1807 (2002).
35. Y. Seki *et al.*, *Development* **134**, 2627 (2007).
36. P. Hajkova *et al.*, *Science* **329**, 78 (2010).
37. H. D. Morgan, W. Dean, H. A. Coker, W. Reik, S. K. Petersen-Mahrt, *J. Biol. Chem.* **279**, 52353 (2004).
38. P. Hajkova *et al.*, *Nature* **452**, 877 (2008).
39. J. Oswald *et al.*, *Curr. Biol.* **10**, 475 (2000).
40. W. Mayer, A. Niveleau, J. Walter, R. Fundele, T. Haaf, *Nature* **403**, 501 (2000).
41. F. Santos, B. Hendrich, W. Reik, W. Dean, *Dev. Biol.* **241**, 172 (2002).
42. T. Nakamura *et al.*, *Nat. Cell Biol.* **9**, 64 (2007).
43. C. R. Farthing *et al.*, *PLoS Genet.* **4**, e1000116 (2008).
44. M. Wossidlo *et al.*, *EMBO J.* **29**, 1877 (2010).
45. Y. Okada, K. Yamagata, K. Hong, T. Wakayama, Y. Zhang, *Nature* **463**, 554 (2010).
46. R. Hirasawa *et al.*, *Genes Dev.* **22**, 1607 (2008).
47. X. Li *et al.*, *Dev. Cell* **15**, 547 (2008).
48. S. Yamanaka, H. M. Blau, *Nature* **465**, 704 (2010).
49. N. Bhutani *et al.*, *Nature* **463**, 1042 (2010).
50. J. Hanna *et al.*, *Nature* **462**, 595 (2009).
51. K. Kim *et al.*, *Nature* **467**, 285 (2010).
52. I. Stancheva, O. El-Maarri, J. Walter, A. Niveleau, R. R. Meehan, *Dev. Biol.* **243**, 155 (2002).
53. R. Öllinger *et al.*, *PLoS Genet.* **4**, e1000199 (2008).
54. F. K. Teixeira *et al.*, *Science* **323**, 1600 (2009); published online 29 January 2009 (10.1126/science.1165313).
55. D. J. Katz, T. M. Edwards, V. Reinke, W. G. Kelly, *Cell* **137**, 308 (2009).
56. U. Brykczynska *et al.*, *Nat. Struct. Mol. Biol.* **17**, 679 (2010).
57. S. S. Hammoud *et al.*, *Nature* **460**, 473 (2009).
58. O. H. Tam *et al.*, *Nature* **453**, 534 (2008).
59. T. Watanabe *et al.*, *Nature* **453**, 539 (2008).
60. S. Ito *et al.*, *Nature* **466**, 1129 (2010).
61. We thank J. A. Law for reading and commenting on the manuscript, and F. Santos for help with figures. S.F. is a Special Fellow of the Leukemia & Lymphoma Society. Supported by NIH grant GM60398 (S.E.J.) and by grants from the UK Biotechnology and Biological Sciences Research Council, UK Medical Research Council, and European Union (W.R.). S.E.J. is an investigator of the Howard Hughes Medical Institute.