NATURE|Vol 447|24 May 2007|doi:10.1038/nature05917

Epigenetic inheritance in plants

Ian R. Henderson[1] & Steven E. Jacobsen[1]

The function of plant genomes depends on chromatin marks such as the methylation of DNA and the post-translational modification of histones. Techniques for studying model plants such as *Arabidopsis thaliana* have enabled researchers to begin to uncover the pathways that establish and maintain chromatin modifications, and genomic studies are allowing the mapping of modifications such as DNA methylation on a genome-wide scale. Small RNAs seem to be important in determining the distribution of chromatin modifications, and RNA might also underlie the complex epigenetic interactions that occur between homologous sequences. Plants use these epigenetic silencing mechanisms extensively to control development and parent-of-origin imprinted gene expression.

Eukaryotic genomes are covalently modified with a diverse set of chromatin marks, which are present on both the DNA and the associated histones (see page 407). Although these changes do not alter the primary DNA sequence, they are frequently heritable through cell division, sometimes for multiple generations, and can thus often be classified as epigenetic marks. These conserved epigenetic marks have been found to influence many aspects of gene expression and chromosome biology, and they have characteristic genomic distributions.

The size of eukaryotic genomes varies extensively and does not correlate with gene number[1]. This is often because of the presence of large amounts of non-gene sequences, which can include pseudogenes, transposable elements, integrated viruses and simple repeats[1]. At the chromosomal level, genomes are organized into euchromatin, which is gene-rich, and heterochromatin, which is repeat-rich[2]. Heterochromatin is defined by three main properties: greater compaction than other genomic regions during interphase, lower accessibility than other regions to transcription and recombination machinery, and the formation of structured nucleosome arrays[2] (see page 399). The defining characteristics of heterochromatin depend on epigenetic information, including post-translational modification of histones and methylation of cytosine bases in DNA[2,3]. The silencing of transposable-element sequences within heterochromatin is probably a genome-defence strategy. However, heterochromatin can also have important roles during chromosomal segregation[4], and transposons and epigenetic silencing have been shown to both modulate gene expression and contribute to *cis*-regulatory sequences[5,6]. Plant systems have been a rich source for the study of epigenetic inheritance, and examples of important discoveries include transposable elements[7], paramutation[8], small interfering RNAs (siRNAs)[9] and RNA-directed DNA methylation[10].

Genomic resources for studying the model plant *Arabidopsis thaliana* have begun to provide insight into the epigenetic 'landscape' of this organism[11,12]. *A. thaliana* has a compact ~130-megabase (Mb) genome, although it contains considerable amounts of heterochromatin, which is repeat-rich and largely located in the centromeric and pericentromeric regions[13,14] (Fig. 1). High-resolution mapping of cytosine methylation by using whole-genome microarrays has confirmed previous reports, showing that this modification co-localizes with repeat sequences and with the centromeric regions[11,12,15]. Fewer than 5% of expressed genes were shown to have methylated promoters, although about one-third of genes were methylated in their open reading frame[11,12]. The significance of methylation in the body of a gene is not fully understood, but such methylation was found to correlate with genes that are both highly transcribed and constitutively expressed[11,12]. By contrast, genes with methylated promoters had lower expression levels and frequently had tissue-specific expression patterns[11,12]. This distribution of cytosine methylation is in contrast to that observed in mammalian genomes, which are often densely methylated but have hypomethylated CG islands in gene promoters[3]. It will be important to describe the 'methylome' of other repeat-rich plant genomes, such as those of the grasses, to test the generality of the patterns observed in *A. thaliana*. Here, we review the emerging and prominent role of RNA in epigenetic inheritance in plants and how such mechanisms are used to control development.

Mediating silencing with RNA

A central question in understanding the epigenetic regulation of genomes is how sequences are recognized or avoided as targets for silencing. There is an increasing appreciation that siRNAs, which are generated by the RNA interference (RNAi) pathway, can provide sequence specificity to guide epigenetic modifications in a diverse range of eukaryotes. Well-studied examples include transcriptional silencing in yeast[16] (see page 399), cytosine methylation in plants[10,17] and genome rearrangements in ciliates[18]. RNA-directed DNA methylation was discovered in tobacco, in which genomic sequences homologous to infectious RNA viroids were found to become cytosine methylated[10]. Subsequently, the expression of double-stranded RNA (dsRNA) in plants was shown to generate siRNAs and cause dense cytosine methylation of homologous DNA in all sequence contexts[19]. This is reflected by the high coincidence of endogenous siRNA clusters with methylated sequences and repeats in *A. thaliana*[11,12,15,20].

All known *de novo* DNA methylation in *A. thaliana* is carried out by DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), which is a homologue of the mammalian DNA methyltransferase 3 (DNMT3) enzymes[21–24] (Fig. 2b). DRM2 can be targeted to a sequence by siRNAs generated from the expression of either direct or inverted repeats[23,24]. Plants encode multiple homologues of the RNAi-machinery components, some of which are specialized for function in RNA-directed DNA methylation[25,26]. The endoribonuclease DICER-LIKE 3 (DCL3) generates 24-nucleotide siRNAs, which are loaded into the PAZ- and PIWI-domain-containing protein ARGONAUTE 4 (AGO4)[26–31] (Fig. 2a). These AGO4-associated siRNAs are proposed to guide the cytosine-methyltransferase activity of DRM2 (refs 26–31). The mechanism by which siRNAs target epigenetic modifications is poorly understood and could involve either DNA–RNA or RNA–RNA hybridization events. Interestingly, epigenetic modifications guided by AGO4 in

[1]Department of Molecular, Cell and Developmental Biology, Howard Hughes Medical Institute, University of California, Los Angeles, California 90095, USA.

A. thaliana have been shown to depend partly on the RNaseH ('slicer') catalytic activity of AGO4 (ref. 30). This could be taken as support for RNA–RNA hybridization having an important role in the targeting of epigenetic modifications.

The accumulation of siRNAs associated with RNA-directed DNA methylation in *A. thaliana* often depends on RNA-DEPENDENT RNA POLYMERASE 2 (RDR2) and the plant-specific protein NUCLEAR RNA POLYMERASE IV A (also known as NUCLEAR RNA POLYMERASE D 1A; NRPD1A), which are involved in a putative amplification pathway[26,32–35] (Fig. 2a). Together, RDR2 and NRPD1A might generate dsRNA substrates for DCL3 to process into siRNAs, although how these proteins are recruited to target loci is unknown. Several loci also show dependence on AGO4 and DRM2 for siRNA accumulation, suggesting that there might be a feedback loop between transcriptional silencing and siRNA generation[24,26].

NRPD1A functions in a complex with NRPD2. A variant of this NRPD complex, which contains NRPD1B instead of NRPD1A, is also required for RNA-directed DNA methylation but participates less frequently in siRNA accumulation[33,35] (Fig. 2a). One possible function for the NRPD1B-containing complex is to generate a target transcript that can hybridize with siRNA-loaded AGO4-containing complexes. Indeed, AGO4 has been observed to bind directly to NRPD1B[28]. The SWI–SNF-family chromatin-remodelling protein DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1) is also required for RNA-directed DNA methylation and could function to facilitate access of DRM2 to target DNA[27,36]. Recently, several proteins in the RNA-directed DNA-methylation pathway have been found to localize to distinct nuclear bodies, including the Cajal body, which is a centre for the processing and modification of many non-coding RNAs[28,29]. Localization to these bodies might be required for the efficient loading of AGO4-containing complexes with siRNA before these complexes travel to the nucleoplasm and, together with DRM2, direct RNA-directed DNA methylation.

Plants show extensive methylation of cytosine bases in the CG, CNG (where N denotes any nucleotide) and CHH (where H denotes A, C or T) sequence contexts[37]. By contrast, most cytosine methylation in mammals is found in the CG sequence context[3,38]. CG methylation is maintained by the homologous proteins METHYLTRANSFERASE 1 (MET1) and DNMT1 in plants and mammals, respectively[39,40] (Fig. 2b). DNMT1 has a catalytic preference for hemimethylated substrates, providing an attractive model for the efficient maintenance of CG methylation after DNA replication and during cell division[38]. Most non-CG methylation in plants is maintained redundantly by DRM2 and the plant-specific protein CHROMOMETHYLASE 3 (CMT3)[23,37] (Fig. 2b); however, some loci show residual non-CG methylation in *drm1 drm2 cmt3* triple mutants, which might be maintained by MET1 (ref. 25). Non-CG methylation differs from CG methylation, because it seems to require an active maintenance signal after DNA replication. At some loci, siRNAs seem to provide this signal, acting through DRM2 activity: for example, at the *MEA-ISR* locus (*MEDEA INTERSTITIAL SUBTELOMERIC REPEATS* locus, an array of seven tandem repeats located downstream of the *MEDEA* gene), the repeats lose all non-CG methylation in *drm2* mutants and in several RNAi-pathway mutants such as *ago4* and *rdr2* (refs 23, 37). By contrast, other loci — for example, the SINE-class retrotransposon *AtSN1* — completely lose non-CG methylation only in *drm1 drm2 cmt3* triple mutants. At *AtSN1*, CMT3 contributes to the maintenance of both CNG methylation and asymmetrical (CHH) methylation. The activity of CMT3 largely depends on the main methyltransferase for H3K9 (the lysine residue at position 9 of histone H3) — SU(VAR)3-9 HOMOLOGUE 4 (SUVH4; also known as KRYPTONITE) — showing that histone methylation is also an important signal for the maintenance of non-CG methylation[41,42]. At present, the factors that determine the relative importance of the RNAi pathway and histone methylation for the maintenance of non-CG methylation at different loci remain unclear.

Communication of silent information

Epigenetically silent expression states can show remarkable stability throughout mitosis and meiosis, although they can retain the ability to revert to an active state[2]. This gives rise to the concept of the epigenetic allele (epiallele), which is defined as an allele that shows a heritable difference in expression as a consequence of epigenetic modifications and not changes in DNA sequence. For example, hypermethylated (silent) epialleles of *SUPERMAN* (which is involved in floral development) known as *clark kent* are stable during many generations of inbreeding, but they can revert to an unmethylated (active) state at a frequency of ~3% per generation[43]. Another notable characteristic of certain epialleles is their ability to influence other homologous sequences both *in cis* and

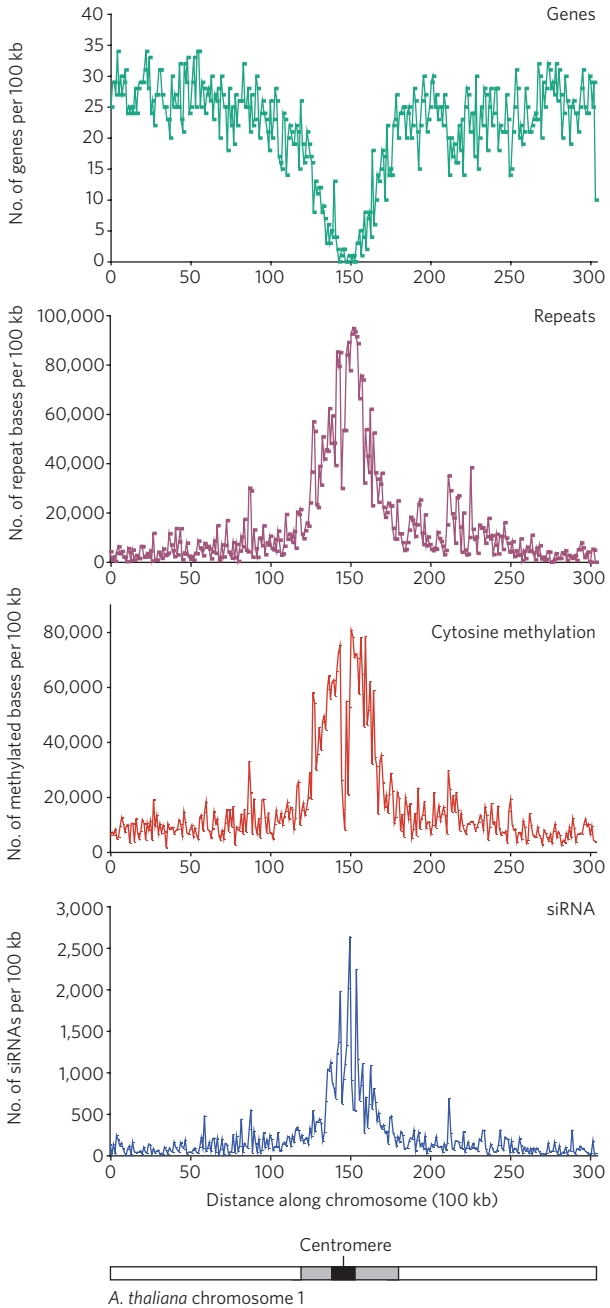

Figure 1 | The epigenetic 'landscape' of *A. thaliana*. The relative abundance of genes (number of annotated genes[11]), repeats (repeat bases per 100 kb; ref. 11), cytosine methylation (methylated bases per 100 kb; ref. 11) and siRNAs (cloned siRNAs per 100 kb; ref. 20) is shown for the length of *A. thaliana* chromosome 1, which is ~30 Mb. Numbers on the *x* axis represent 100-kb windows along the chromosome. A diagram of chromosome 1 is also shown, with white bars indicating euchromatic arms, grey bars indicating pericentromeric heterochromatin and the black bar indicating the centromeric core. (Figure courtesy of X. Zhang, University of California, Los Angeles.)

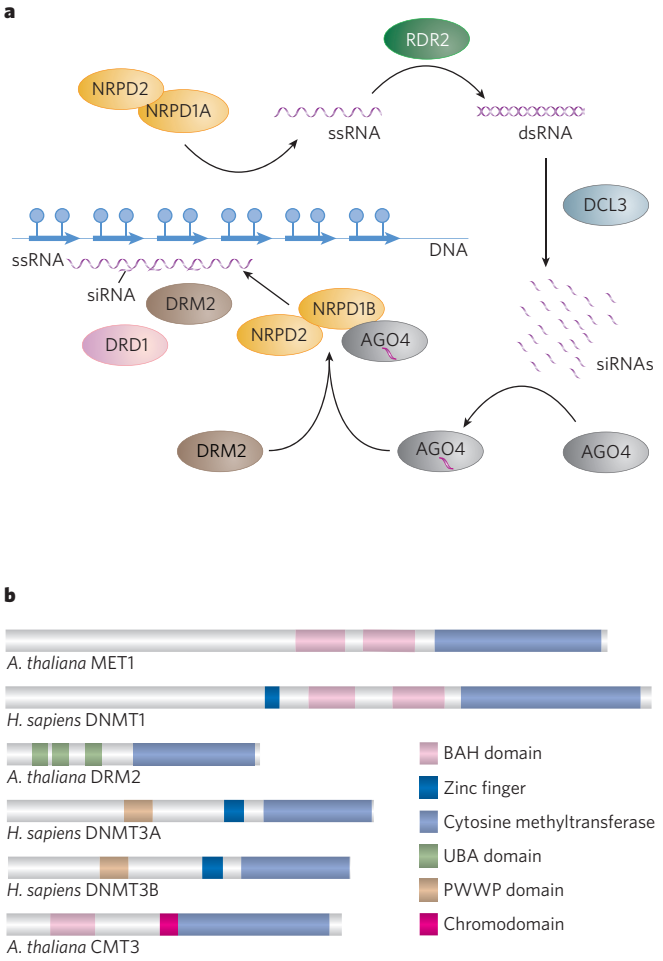

Figure 2 | RNA-directed DNA methylation. a, Putative pathway for RNA-directed DNA methylation in *A. thaliana*. Target loci (in this case tandemly repeated sequences; coloured arrows) recruit an RNA polymerase IV complex consisting of NRPD1A and NRPD2 through an unknown mechanism, and this results in the generation of a single-stranded RNA (ssRNA) species. This ssRNA is converted to double-stranded RNA (dsRNA) by the RNA-dependent RNA polymerase RDR2. The dsRNA is then processed into 24-nucleotide siRNAs by DCL3. The siRNAs are subsequently loaded into the PAZ- and PIWI-domain-containing protein AGO4, which associates with another form of the RNA polymerase IV complex, NRPD1B–NRPD2. AGO4 that is 'programmed' with siRNAs can then locate homologous genomic sequences and guide the protein DRM2, which has *de novo* cytosine methyltransferase activity. Targeting of DRM2 to DNA sequences also involves the SWI–SNF-family chromatin-remodelling protein DRD1. The NRPD1B–NRPD2 complex might generate a target transcript (ssRNA) to which the AGO4-associated siRNAs can hybridize. Given that siRNAs homologous to some loci are absent in *drm2* mutants and *ago4* mutants, it is possible that DNA methylation (blue circles) also stimulates siRNA generation and reinforces silencing. **b**, DNA methyltransferase structure and function. Plant and mammalian genomes encode homologous cytosine methyltransferases, of which there are three classes in plants and two in mammals. *A. thaliana* MET1 and *Homo sapiens* (human) DNMT1 both function to maintain CG methylation after DNA replication, through a preference for hemimethylated substrates, and both have amino-terminal bromo-adjacent homology (BAH) domains of unknown function. *De novo* DNA methylation is carried out by the homologous proteins DRM2 (in *A. thaliana*) and DNMT3A and DNMT3B (both in *H. sapiens*). Despite their homology, these proteins have distinct N-terminal domains, and the catalytic motifs present in the cytosine methyltransferase domain are ordered differently in DRM2 and the DNMT3 proteins. Plants also have another class of methyltransferase, which is not found in mammals. CMT3 functions together with DRM2 to maintain non-CG methylation. PWWP, Pro-Trp-Trp-Pro motif; UBA, ubiquitin associated.

in trans[2]. One example is paramutation, which was discovered in plants and is defined as allelic interactions that cause a meiotically heritable change in the expression of one of the alleles[8]. *Trans*-phenomena similar to paramutation have also been described in mammals, including at a chimaeric version of the mouse *Rasgrf1* (Ras protein-specific guanine-nucleotide-releasing factor 1) locus that contained the imprinting control region from the insulin-like growth factor 2 receptor gene[44].

One of the best-studied paramutation systems is the maize (*Zea mays*) locus *b1*, which encodes a transcription factor that is required for accumulation of the pigment anthocyanin[8]. The paramutagenic epiallele *B′*, which causes light pigmentation, arises spontaneously at a low frequency from its paramutable parent allele *B-I*, which causes dark pigmentation[45]. *B′* epialleles convert *B-I* alleles to *B′* epialleles when heterozygous with 100% penetrance, and the newly created paramutated *B′* epialleles can pass on their silent state in subsequent crosses[45] (Fig. 3). *B′* epialleles are transcribed at one-twentieth to one-tenth the rate of *B-I* alleles but have identical gene sequences[45,46]. Fine-structure recombination mapping of alleles resulting from a cross between individuals with paramutagenic alleles and those with neutral alleles (which cannot participate in paramutation) enabled the sequences required for paramutation to be defined; these sequences are present as an array of 7 tandem 853-base repeats, which is located ~100 kilobases (kb) upstream of *b1* (refs 45, 46). The sequences are present as a single copy in neutral alleles. Recombinant alleles with three repeats show partial paramutational ability, whereas alleles with seven repeats are fully active in paramutation[45,46]. These repeats were also shown to have a closed chromatin structure and more cytosine methylation in *B′* epialleles than in *B-I* alleles[46]. However, for *B′*, cytosine methylation was found to be established after the silent state, so it is unlikely to be the cause[46]. There are several models of *trans*-communication between alleles, including physical pairing of alleles and transmission of an RNA signal. A model for paramutagenic interactions being mediated by siRNA is supported by the finding that a genetic suppressor of paramutation, *mediator of paramutation1* (*mop1*), encodes the maize orthologue of the RNA-dependent RNA polymerase RDR2 (refs 47, 48). So far, siRNAs homologous to the tandem repeats upstream of *B′* have not been reported, although such repeats are commonly associated with small RNAs[20,49]. The *mop1* gene is also required for silencing transgenes and *Mutator*-like transposons, indicating that RNA-dependent RNA polymerases and siRNAs have a role in heterochromatic silencing in monocotyledonous plants[50]. The detailed relationships between siRNAs, chromatin structure at the repeats upstream of *B′*, and the ability to transfer epigenetic states will be intriguing to determine.

The *A. thaliana* gene *FWA* has similarities to maize *b1* in that it has tandem repeats upstream that, when methylated, cause heritable silencing of expression[51]. Stably hypomethylated *fwa-1* epialleles have been found to be generated spontaneously and in *met1* mutant backgrounds[39,40,51], causing overexpression of the transcription factor FWA and a dominant late-flowering phenotype[51]. In contrast to *B′* epialleles, methylated and unmethylated *fwa* epialleles are not influenced by the presence of one another in heterozygotes[23,49,51]. However, introduction of unmethylated transgenic copies of *FWA* by *Agrobacterium tumefaciens*-mediated transformation leads to efficient *de novo* silencing of the incoming transgene, in a process that depends on both DRM2 and the RNA-directed DNA-methylation RNAi pathway[22,23] (Fig. 3). Intriguingly, an unmethylated *FWA* transgene obtained after transformation into a *drm2* mutant does not become remethylated after outcrossing to wild-type *A. thaliana*[22,23]. This finding suggests that, during the transformation process, there is a 'surveillance' window when the incoming *FWA* transgene is competent to be silenced. *A. tumefaciens* targets the female gametophyte (which is haploid) during transformation, but introduction of *FWA* into *DRM2/drm2* heterozygotes revealed that the silencing window must be present after fertilization[49]. Structure–function analysis of an *FWA* transgene showed that the upstream tandem repeats are necessary and sufficient for transformation-dependent silencing and were also found to produce homologous siRNA[49]. Interestingly, the efficiency by which an incoming

FWA transgene is silenced can be influenced by the methylation state of endogenous *FWA*[49]. Whereas introduction of an *FWA* transgene into a background in which the endogenous *FWA* gene is methylated leads to extremely efficient silencing of the transgene, transformation into the *fwa-1* background, which contains an unmethylated endogenous gene, leads to inefficient methylation and silencing of the *FWA* transgene[49] (Fig. 3). Furthermore, an introduced transgene can occasionally cause silencing of the unmethylated *fwa-1* endogenous gene[49]. These results reveal extensive communication between the transgenic and endogenous *FWA* gene copies during transformation, and this communication depends on the DNA methylation state of the endogenous gene. Surprisingly, these differences between *fwa-1* epialleles are not accounted for by siRNA production, because the repeat-derived siRNAs accumulate equally in plants with wild-type *FWA* and those with *fwa-1* (ref. 49). Hence, recruitment of siRNA machinery to a locus is not always sufficient for RNA-directed DNA methylation and probably also requires modifications of chromatin.

Maintenance of silencing at *FWA* depends mainly on CG methylation, because *met1* alleles generate hypomethylated *fwa-1* epialleles at a high frequency[39,40]. Although the tandem repeats upstream of *FWA* are also methylated at non-CG sequences, loss of this methylation in *drm1 drm2 cmt3* triple mutants does not cause reactivation and late flowering[37]. Genome-wide analysis of cytosine methylation and transcription in *drm1 drm2 cmt3* triple mutants has identified genes with methylated promoters, the expression of which depends strongly on DRM- and CMT3-mediated non-CG methylation[11]. These methylated genes might be responsible for the developmental phenotypes of *drm1 drm2 cmt3* triple mutants, which include misshapen leaves and reduced stature[27,37].

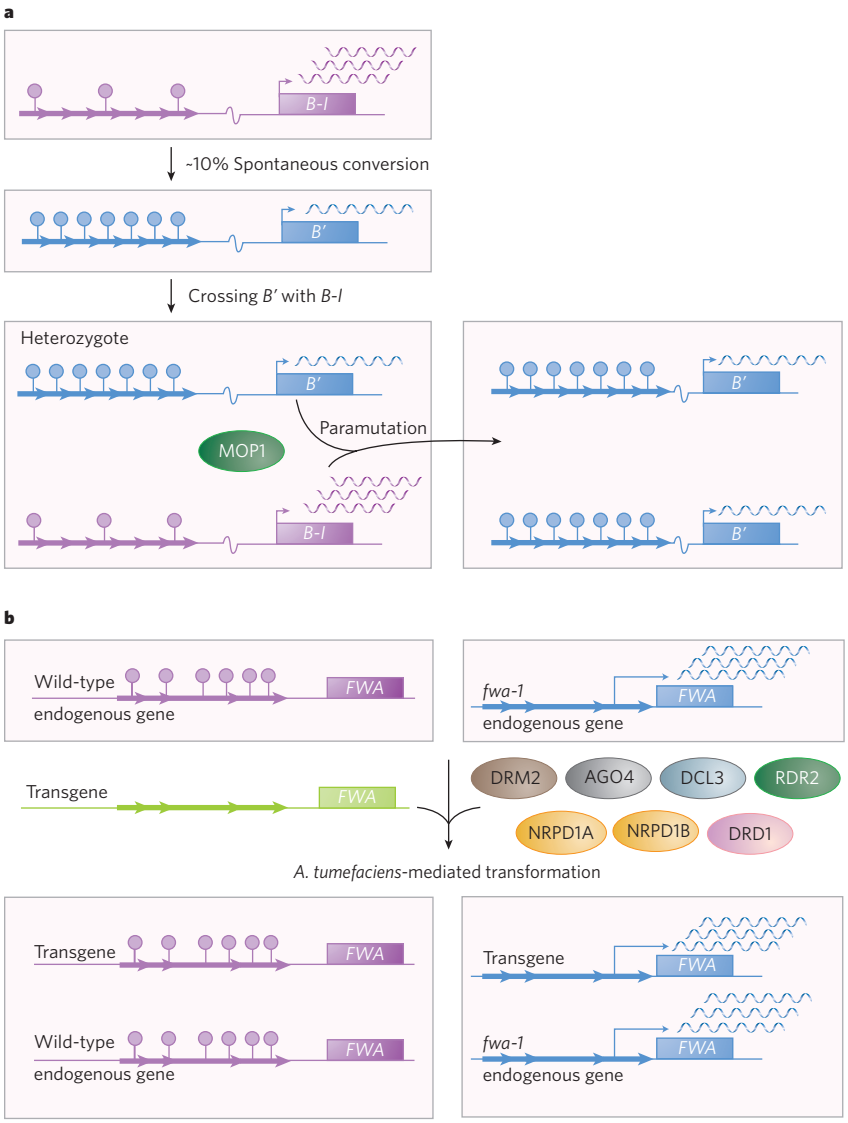

Figure 3 | *Trans*-epiallele interactions at *b1* and *FWA*. **a**, Paramutation at the *b1* locus in maize. The *B-I* allele (pink) of the *b1* gene in maize has an upstream tandem-repeat region (coloured arrows) and spontaneously gives rise to silenced *B′* epialleles (blue) at a low frequency. *B′* epialleles are more heavily methylated at cytosine bases in the repeat region and are less frequently transcribed. When the *B′* epiallele is brought together with a new copy of *B-I* by crossing of maize plants, the *B-I* allele is paramutated to a silenced *B′* state with 100% penetrance. *Trans*-communication between epialleles requires MOP1, the maize homologue of *A. thaliana* RDR2, suggesting that siRNA-mediated silencing might be involved in the conversion of *B-I* to *B′*. **b**, *De novo* silencing of *FWA* transgenes in wild-type and *fwa-1 A. thaliana*. The *FWA* gene in wild-type *A. thaliana* (pink) is methylated at cytosine bases in a pair of tandem repeats in its promoter, silencing its expression. Mutations that decrease DNA methylation give rise to hypomethylated *fwa-1* epialleles (blue), which overexpress the transcription factor FWA, thereby causing late flowering. Introduction of an unmethylated *FWA* transgene (green) by *A. tumefaciens*-mediated transformation of wild-type plants results in efficient methylation and silencing of the incoming transgene. This process depends on DRM2, AGO4, DCL3, RDR2, NRPD1A, NRPD1B and DRD1. By contrast, transformation of an *fwa-1* background results in inefficient silencing of the transgene, indicating that the methylation state of endogenous *FWA* is important for transgene silencing.

In contrast to the independently segregating epialleles that arise in *met1* mutants (as a result of the stable loss of CG methylation)[39,40,51], backcrossing *drm1 drm2 cmt3* triple mutants to wild-type plants or reintroducing either *DRM2* or *CMT3* by transformation immediately rescues these morphological phenotypes[27]. This finding suggests that non-CG methylation can be more easily re-established, possibly allowing flexible regulation of genes. However, it is unclear how commonly this type of regulation is used, because few examples of DNA-methylation-regulated plant genes have been described.

Silencing through time and development

The life cycles of plants differ from those of animals in that the products of meiosis undergo mitotic proliferation to form multicellular gametophytes (that is, the embryo sac and the pollen in flowering plants). The embryo sac (female) contains an egg cell, which is haploid, and this is fertilized by a sperm nucleus, which is also haploid, to form a diploid embryo. A second sperm nucleus fertilizes the central cell, which is diploid, to form triploid endosperm, an extra-embryonic tissue that has a supportive role during embryogenesis. The central cell and the endosperm show parent-of-origin-dependent monoallelic expression, or imprinting, which is important for proper seed development[52]. For example, in *A. thaliana*, the tandem repeats of maternal *FWA* alleles are specifically demethylated in the central cell and the endosperm, leading to expression of *FWA* in these tissues[53]. Demethylation and activation of *FWA* depend on maternal expression of the gene encoding the DNA glycosylase–lyase DEMETER (DME), which can directly excise the base 5-methylcytosine[54–56]. Because the endosperm is a terminally differentiating extra-embryonic tissue, this mechanism does not necessitate remethylation of *FWA*[53]. This is in contrast to mammals, in which demethylation of imprinted genes occurs in primordial germ cells (the cells that ultimately generate the germ line) and is followed by germline-specific remethylation and silencing (see page 425). Other imprinted genes such as *MEA* and *FERTILIZATION-INDEPENDENT SEED 2* also have cytosine-methylated regions in their promoters that are associated with maternally restricted expression[55,57]. However, only for *FWA* has it been shown that differential methylation of particular sequences is required for the regulation of imprinting[53,58].

Cytosine demethylation is also likely to have an important role in the control of silencing in situations other than gametophytic generation and imprinting. *DME* belongs to a small *A. thaliana* gene family that includes the somatically expressed gene *REPRESSOR OF SILENCING 1* (*ROS1*)[54,59]. Mutations in *ROS1* have been shown to increase RNA-directed DNA methylation, and ROS1 has been shown to function as a cytosine demethylase[56,59,60]. Together, these exciting discoveries have defined a long-sought cytosine demethylation pathway, and they raise many interesting questions. For example, to what extent are genomic methylation patterns balanced by the targeting of *de novo* DNA methyltransferases and DNA glycosylases? Furthermore, there are indications of a similar mechanism for cytosine demethylation in vertebrates[61,62].

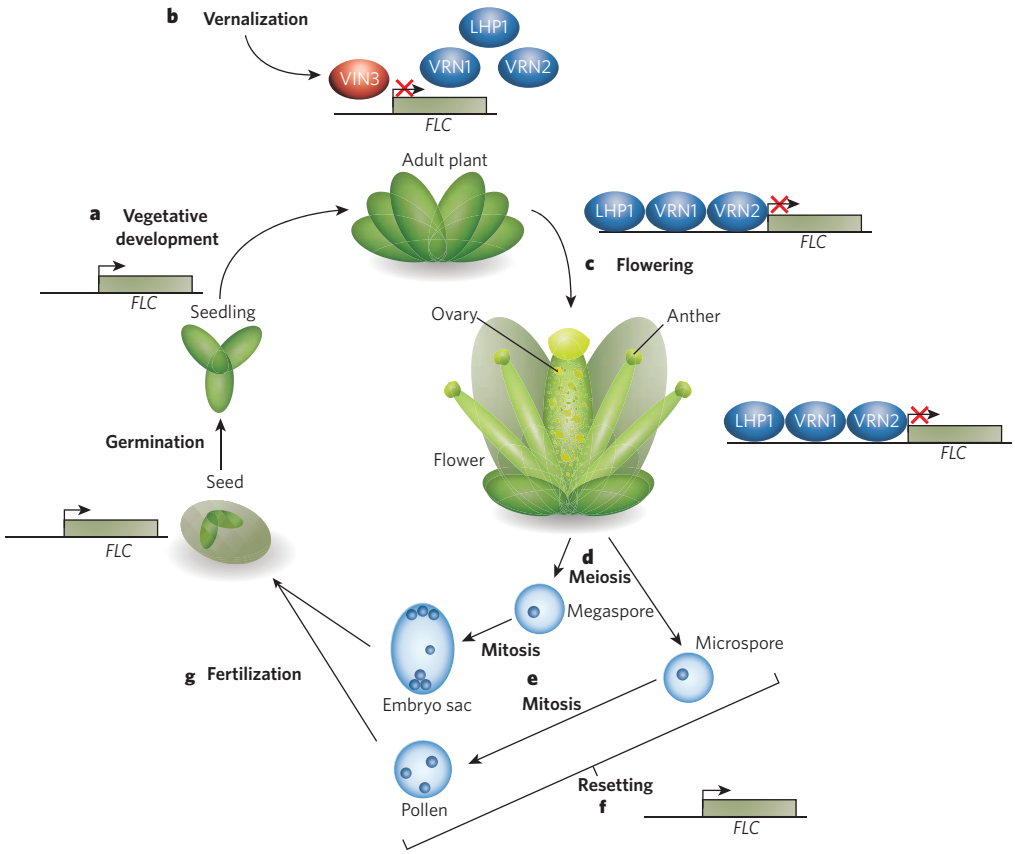

Figure 4 | PcG-protein-mediated silencing throughout the *A. thaliana* life cycle. The activation state of the PcG protein target *FLC* is illustrated throughout the plant life cycle. **a**, *FLC* is transcriptionally active in seeds and seedlings, preventing the plant from flowering and prolonging vegetative development. **b**, Exposure to a long period of cold (that is, vernalization) results in the expression of VIN3 (red), which initiates repression of *FLC* transcription, and the binding of the PcG protein VRN2, as well as VRN1 and LHP1 (blue). In this process, chromatin at *FLC* is epigenetically modified by the trimethylation of H3K27. **c**, After warmer temperatures return, *FLC* repression is maintained, allowing flowering to be induced by other cues. **d**, During flower development, the anthers and ovaries are sites of meiotic differentiation, giving rise to haploid cells known as microspores and megaspores, respectively. **e**, These meiotic products undergo mitotic proliferation to form the multicellular embryo sac and pollen gametophytes. **f**, PcG-protein-mediated repression at *FLC* is removed during an undefined resetting process. **g**, Then, the pollen contributes sperm nuclei to the embryo sac, and these fertilize the haploid egg cell and diploid central cell (not shown), forming the embryo and endosperm (respectively) in a new seed, in which *FLC* is re-expressed.

Other examples of imprinted genes are maize *fertilization-independent endosperm1* (*fie1*) and *fie2*, which show monoallelic expression from the maternal allele during endosperm development. This is reflected by the promoters of the silent paternal alleles having differentially methylated regions (DMRs)[63,64]. Analysis of DMR methylation of *fie* alleles in sperm, egg and central cells showed interesting differences in the mechanism for imprinting *fie1* and *fie2* (ref. 64). The DMR of *fie1* is heavily methylated in all three cell types, but the maternal alleles in the central cell (which contribute to the endosperm) become specifically demethylated, resembling the imprinting mechanism described for *A. thaliana FWA*[64]. By contrast, the DMR of *fie2* is unmethylated in all gametes, although the paternal allele becomes methylated *de novo* in the endosperm. Furthermore, the *fie2* DMR also showed extensive non-CG methylation, which is consistent with a DRM2-type-mediated RNA-directed DNA methylation process[64]. A further instance of potential gene regulation by *de novo* DNA methylation is provided by the *Brassica rapa SP11* locus, which encodes a pollen self-incompatibility determinant[65]. The *B. rapa* self-incompatibility phenotype is controlled by dominance relationships between *S*-haplotypes, and recessive *SP11* alleles were found to be specifically methylated *de novo* and silenced in the anther tapetal tissues[65]. It will be interesting to determine the prevalence of such instances of tissue-specific gene regulation by DNA methylation.

In addition to the gametophytic tissues being an important location for the establishment of imprinted gene expression, they also maintain pre-existing patterns of cytosine methylation. Evidence that silencing is important during gametophytic generation is provided by null *met1* alleles in *A. thaliana*, which produce hypomethylated epialleles even when the individual is heterozygous for the null allele[40]. This is caused by loss of cytosine methylation in the gametophytes of *met1* mutants, a loss that is greater when *met1* is inherited through the female gametophyte than the male[40]. This difference is probably accounted for by the female gametophyte (that is, the embryo sac) undergoing one more postmeiotic round of DNA replication before fertilization than the male gametophyte (that is, the pollen)[40].

A different epigenetic system used to developmentally silence genes during plant life cycles involves Polycomb group (PcG) proteins[66]. A conserved complex known as Polycomb repressive complex 2 (PRC2) functions to maintain patterns of gene repression in both plants and animals, using H3K27 methylation[66] (see page 425). However, in plants, there are several PRC2 complexes, with overlapping subunit compositions, specialized for distinct developmental roles[66]. For example, the PcG proteins have an important role in the regulation of imprinted gene expression. *A. thaliana MEA*, which is a homologue of *Drosophila melanogaster Enhancer of zeste*, shows maternally imprinted expression[67]. An important component of *MEA* imprinting is repression of the paternal *MEA* allele in the endosperm, and this process has been found to involve *MEA* autoregulation, using H3K27 trimethylation[55,68,69]. Interestingly, the mammalian PcG protein EED (embryonic ectoderm development) has also been shown to have an important role in the control of imprinted gene expression[70].

Another well-understood example of PcG-protein-mediated regulation in plants involves silencing of the floral-repressor gene *FLOWERING LOCUS C* (*FLC*) during the vernalization response in *A. thaliana*[71–73] (Fig. 4). Expression of *FLC*, which encodes a MADS-box-containing transcription factor, delays flowering and can be silenced by exposure of the plant to long periods of cold (that is, vernalization)[71–73]. In nature, this cold treatment occurs in winter and leads to flowering in favourable spring conditions. After the cold signal has been removed, *FLC* silencing is stable[71–73]. Mutations in the *VERNALIZATION 2* (*VRN2*) gene, which encodes a homologue of the *D. melanogaster* PcG protein Suppressor of zeste 12, cause late flowering after vernalization as a result of high levels of *FLC* expression[72]. Interestingly, *vrn2* mutants can silence *FLC* expression during the cold but fail to maintain this repression after the cold signal has been removed[72]. VRN2 is also required for acquisition of H3K27 dimethylation and trimethylation at *FLC* during vernalization, consistent with the known functions of PRC2 in maintaining patterns of gene repression[71,73,74].

The mechanism by which the vernalization-specific PcG-protein complex is recruited to *FLC* is not well understood but is known to require the PHD-finger-domain-containing protein VERNALIZATION INSENSITIVE 3 (VIN3)[73]. Because *VIN3* expression is induced after cold treatment, this protein might be a component of the signalling pathway that recruits PcG-protein-mediated repression to *FLC*[73] (Fig. 4). Recently, the *A. thaliana* homologue of *D. melanogaster* Heterochromatin protein 1 (HP1) — LIKE HETEROCHROMATIN PROTEIN 1 (LHP1; also known as TFL2) — was found to be required for the maintenance of *FLC* silencing after vernalization[75,76]. LHP1 becomes associated with the silenced *FLC* locus, a process that depends on an intronic sequence element[76]. The role of LHP1 in the repression of PcG-protein-regulated genes differs markedly from the main role of animal HP1 in heterochromatic silencing (see page 399). The DNA-binding protein VRN1 is also required for the maintenance of *FLC* silencing and associates with mitotic chromosomes[75,77]. Interestingly, VRN1 is absent from meiotic chromosomes of developing pollen[75]. One speculation is that this absence is associated with the resetting of *FLC* expression, which leads to a requirement for vernalization, at the start of each generation. Indeed, all PcG-protein-mediated silencing might be reset at some point during meiosis or gametogenesis, through an unknown mechanism (Fig. 4).

Conclusions

Plants continue to be excellent systems for the study of epigenetics, and their silencing mechanisms have marked similarities to those of mammals. An advantage of using plants is that they are tolerant of genome stresses, such as large losses of DNA methylation and changes in chromosome number. The elegant genetic tools available for organisms such as maize and *A. thaliana* are facilitating the dissection of epigenetic control. Recent advances such as the development of whole-genome microarrays and high-throughput sequencing are allowing the generation of large-scale data sets for epigenetic modifications and small RNAs that are extending our view to a genome-wide scale. Together, these approaches should enable major advances in our understanding of epigenetics to be made using plant systems: for example, how specific chromatin modifications are established and maintained, how they influence one another, and the extent to which they are used throughout the genome. This work should provide important insight for fields as diverse as cancer biology, development and evolution. ■

1. Gregory, T. R. The C-value enigma in plants and animals: a review of parallels and an appeal for partnership. *Ann. Bot. (Lond.)* **95,** 133–146 (2005).
2. Hall, I. M. & Grewal, S. I. in *RNAi: A Guide to Gene Silencing* (ed. Hannon, G. J.) 205–232 (Cold Spring Harbor Laboratory Press, Woodbury, 2003).
3. Bernstein, B. E., Meissner, A. & Lander, E. S. The epigenome. *Cell* **128,** 669–681 (2006).
4. Bernard, P. *et al.* Requirement of heterochromatin for cohesion at centromeres. *Science* **294,** 2539–2542 (2001).
5. Bejerano, G. *et al.* A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* **441,** 87–90 (2006).
6. Liu, J., He, Y., Amasino, R. & Chen, X. siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in *Arabidopsis*. *Genes Dev.* **18,** 2873–2878 (2004).
7. Comfort, N. C. From controlling elements to transposons: Barbara McClintock and the Nobel Prize. *Trends Biochem. Sci.* **26,** 454–457 (2001).
8. Chandler, V. L. & Stam, M. Chromatin conversations: mechanisms and implications of paramutation. *Nature Rev. Genet.* **5,** 532–544 (2004).
9. Hamilton, A. J. & Baulcombe, D. C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286,** 950–952 (1999).
10. Wassenegger, M., Heimes, S., Riedel, L. & Sanger, H. L. RNA-directed *de novo* methylation of genomic sequences in plants. *Cell* **76,** 567–576 (1994).
11. Zhang, X. *et al.* Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* **126,** 1189–1201 (2006).
12. Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T. & Henikoff, S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature Genet.* **39,** 61–69 (2007).
13. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408,** 796–815 (2000).
14. Fransz, P. F. *et al.* High-resolution physical mapping in *Arabidopsis thaliana* and tomato by fluorescence *in situ* hybridization to extended DNA fibres. *Plant J.* **9,** 421–430 (1996).
15. Lippman, Z. *et al.* Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430,** 471–476 (2004).
16. Volpe, T. A. *et al.* Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297,** 1833–1837 (2002).
17. Aufsatz, W., Mette, M. F., van der Winden, J., Matzke, A. J. & Matzke, M. RNA-directed DNA methylation in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **99** (suppl. 4), 16499–16506 (2002).

18. Mochizuki, K., Fine, N. A., Fujisawa, T. & Gorovsky, M. A. Analysis of a *piwi*-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*. *Cell* **110,** 689–699 (2002).
19. Matzke, M., Matzke, A. J. & Kooter, J. M. RNA: guiding gene silencing. *Science* **293,** 1080–1083 (2001).
20. Lu, C. *et al.* Elucidation of the small RNA component of the transcriptome. *Science* **309,** 1567–1569 (2005).
21. Cao, X. *et al.* Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. *Curr. Biol.* **13,** 2212–2217 (2003).
22. Cao, X. & Jacobsen, S. E. Role of the *Arabidopsis* DRM methyltransferases in *de novo* DNA methylation and gene silencing. *Curr. Biol.* **12,** 1138–1144 (2002).
23. Chan, S. W. *et al.* RNA silencing genes control *de novo* DNA methylation. *Science* **303,** 1336 (2004).
24. Zilberman, D. *et al.* Role of *Arabidopsis ARGONAUTE4* in RNA-directed DNA methylation triggered by inverted repeats. *Curr. Biol.* **14,** 1214–1220 (2004).
25. Henderson, I. R. *et al.* Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nature Genet.* **38,** 721–725 (2006).
26. Xie, Z. *et al.* Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* **2,** e104 (2004).
27. Chan, S. W. *et al.* RNAi, DRD1, and histone methylation actively target developmentally important non-CG DNA methylation in *Arabidopsis*. *PLoS Genet.* **2,** e83 (2006).
28. Li, C. F. *et al.* An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* **126,** 93–106 (2006).
29. Pontes, O. *et al.* The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. *Cell* **126,** 79–92 (2006).
30. Qi, Y. *et al.* Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* **443,** 1008–1012 (2006).
31. Zilberman, D., Cao, X. & Jacobsen, S. E. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299,** 716–719 (2003).
32. Herr, A. J., Jensen, M. B., Dalmay, T. & Baulcombe, D. C. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308,** 118–120 (2005).
33. Kanno, T. *et al.* Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nature Genet.* **37,** 761–765 (2005).
34. Onodera, Y. *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120,** 613–622 (2005).
35. Pontier, D. *et al.* Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* **19,** 2030–2040 (2005).
36. Kanno, T. *et al.* Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr. Biol.* **14,** 801–805 (2004).
37. Cao, X. & Jacobsen, S. E. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc. Natl Acad. Sci. USA* **99** (suppl. 4), 16491–16498 (2002).
38. Goll, M. G. & Bestor, T. H. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74,** 481–514 (2005).
39. Kankel, M. W. *et al. Arabidopsis* MET1 cytosine methyltransferase mutants. *Genetics* **163,** 1109–1122 (2003).
40. Saze, H., Mittelsten Scheid, O. & Paszkowski, J. Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nature Genet.* **34,** 65–69 (2003).
41. Jackson, J. P., Lindroth, A. M., Cao, X. & Jacobsen, S. E. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* **416,** 556–560 (2002).
42. Malagnac, F., Bartee, L. & Bender, J. An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J.* **21,** 6842–6852 (2002).
43. Jacobsen, S. E. & Meyerowitz, E. M. Hypermethylated *SUPERMAN* epigenetic alleles in *Arabidopsis*. *Science* **277,** 1100–1103 (1997).
44. Herman, H. *et al. Trans* allele methylation and paramutation-like effects in mice. *Nature Genet.* **34,** 199–202 (2003).
45. Stam, M. *et al.* The regulatory regions required for *B'* paramutation and expression are located far upstream of the maize *b1* transcribed sequences. *Genetics* **162,** 917–930 (2002).
46. Stam, M., Belele, C., Dorweiler, J. E. & Chandler, V. L. Differential chromatin structure within a tandem array 100 kb upstream of the maize *b1* locus is associated with paramutation. *Genes Dev.* **16,** 1906–1918 (2002).
47. Alleman, M. *et al.* An RNA-dependent RNA polymerase is required for paramutation in maize. *Nature* **442,** 295–298 (2006).
48. Woodhouse, M. R., Freeling, M. & Lisch, D. Initiation, establishment, and maintenance of heritable *MuDR* transposon silencing in maize are mediated by distinct factors. *PLoS Biol.* **4,** e339 (2006).
49. Chan, S. W.-L., Zhang, X., Bernatavichute, Y. V. & Jacobsen, S. E. Two-step recruitment of RNA-directed DNA methylation to tandem repeats. *PLoS Biol.* **4,** e363 (2006).
50. Lisch, D., Carey, C. C., Dorweiler, J. E. & Chandler, V. L. A mutation that prevents paramutation in maize also reverses *Mutator* transposon methylation and silencing. *Proc. Natl Acad. Sci. USA* **99,** 6130–6135 (2002).
51. Soppe, W. J. *et al.* The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol. Cell* **6,** 791–802 (2000).
52. Gehring, M., Choi, Y. & Fischer, R. L. Imprinting and seed development. *Plant Cell* **16,** S203–S213 (2004).
53. Kinoshita, T. *et al.* One-way control of *FWA* imprinting in *Arabidopsis* endosperm by DNA methylation. *Science* **303,** 521–523 (2004).
54. Choi, Y. *et al.* DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell* **110,** 33–42 (2002).
55. Gehring, M. *et al.* DEMETER DNA glycosylase establishes *MEDEA* polycomb gene self-imprinting by allele-specific demethylation. *Cell* **124,** 495–506 (2006).
56. Morales-Ruiz, T. *et al.* DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc. Natl Acad. Sci.USA* **103,** 6853–6858 (2006).
57. Jullien, P. E., Kinoshita, T., Ohad, N. & Berger, F. Maintenance of DNA methylation during the *Arabidopsis* life cycle is essential for parental imprinting. *Plant Cell* **18,** 1360–1372 (2006).
58. Kinoshita, Y. *et al.* Control of *FWA* gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. *Plant J.* **49,** 38–45 (2007).
59. Gong, Z. *et al.* ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell* **111,** 803–814 (2002).
60. Agius, F., Kapoor, A. & Zhu, J. K. Role of the *Arabidopsis* DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc. Natl Acad. Sci. USA* **103,** 11796–11801 (2006).
61. Barreto, G. *et al.* Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature* **445,** 671–675 (2007).
62. Jost, J. P., Siegmann, M., Sun, L. & Leung, R. Mechanisms of DNA demethylation in chicken embryos. Purification and properties of a 5-methylcytosine-DNA glycosylase. *J. Biol. Chem.* **270,** 9734–9739 (1995).
63. Danilevskaya, O. N. *et al.* Duplicated *fie* genes in maize: expression pattern and imprinting suggest distinct functions. *Plant Cell* **15,** 425–438 (2003).
64. Gutierrez-Marcos, J. F. *et al.* Epigenetic asymmetry of imprinted genes in plant gametes. *Nature Genet.* **38,** 876–878 (2006).
65. Shiba, H. *et al.* Dominance relationships between self-incompatibility alleles controlled by DNA methylation. *Nature Genet.* **38,** 297–299 (2006).
66. Kohler, C. & Grossniklaus, U. Epigenetic inheritance of expression states in plant development: the role of Polycomb group proteins. *Curr. Opin. Cell Biol.* **14,** 773–779 (2002).
67. Kinoshita, T., Yadegari, R., Harada, J. J., Goldberg, R. B. & Fischer, R. L. Imprinting of the *MEDEA* polycomb gene in the *Arabidopsis* endosperm. *Plant Cell* **11,** 1945–1952 (1999).
68. Baroux, C., Gagliardini, V., Page, D. R. & Grossniklaus, U. Dynamic regulatory interactions of Polycomb group genes: *MEDEA* autoregulation is required for imprinted gene expression in *Arabidopsis*. *Genes Dev.* **20,** 1081–1086 (2006).
69. Jullien, P. E., Katz, A., Oliva, M., Ohad, N. & Berger, F. Polycomb group complexes self-regulate imprinting of the Polycomb group gene *MEDEA* in *Arabidopsis*. *Curr. Biol.* **16,** 486–492 (2006).
70. Mager, J., Montgomery, N. D., de Villena, F. P. & Magnuson, T. Genome imprinting regulated by the mouse Polycomb group protein Eed. *Nature Genet.* **33,** 502–507 (2003).
71. Bastow, R. *et al.* Vernalization requires epigenetic silencing of *FLC* by histone methylation. *Nature* **427,** 164–167 (2004).
72. Gendall, A. R., Levy, Y. Y., Wilson, A. & Dean, C. The *VERNALIZATION 2* gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. *Cell* **107,** 525–535 (2001).
73. Sung, S. & Amasino, R. M. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. *Nature* **427,** 159–164 (2004).
74. Sung, S., Schmitz, R. J. & Amasino, R. M. A PHD finger protein involved in both the vernalization and photoperiod pathways in *Arabidopsis*. *Genes Dev.* **20,** 3244–3248 (2006).
75. Mylne, J. S. *et al.* LHP1, the *Arabidopsis* homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of *FLC*. *Proc. Natl Acad. Sci. USA* **103,** 5012–5017 (2006).
76. Sung, S. *et al.* Epigenetic maintenance of the vernalized state in *Arabidopsis thaliana* requires LIKE HETEROCHROMATIN PROTEIN 1. *Nature Genet.* **38,** 706–710 (2006).
77. Levy, Y. Y., Mesnage, S., Mylne, J. S., Gendall, A. R. & Dean, C. Multiple roles of *Arabidopsis* VRN1 in vernalization and flowering time control. *Science* **297,** 243–246 (2002).

Acknowledgements We thank S. Chan, C. Fei Li, K. Niakan, M. Ong and all members of the Jacobsen laboratory for useful comments and discussion. We apologize to colleagues whose research we did not have space to discuss. I.R.H. was supported by a long-term fellowship from the European Molecular Biology Organization, a Special Fellow grant from The Leukemia & Lymphoma Society, and a grant from the National Institutes of Health. S.E.J is an investigator of the Howard Hughes Medical Institute.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence should be addressed to S.E.J. (jacobsen@ucla.edu).