Published in final edited form as:

Nat Plants. ; 2(5): 16058. doi:10.1038/nplants.2016.58.

Unique cell-type specific patterns of DNA methylation in the root meristem

Taiji Kawakatsu^{1,2,3,9}, Tim Stuart^{4,9}, Manuel Valdes^{5,9}, Natalie Breakfield⁵, Robert J. Schmitz^{1,2,6}, Joseph R. Nery², Mark A. Urich², Xinwei Han⁵, Ryan Lister^{2,4,*}, Philip N. Benfey^{5,7,*}, and Joseph R. Ecker^{1,2,8,*}

¹Plant Biology Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

²Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

³Genetically Modified Organism Research Center, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan

⁴ARC Centre of Excellence in Plant Energy Biology, The University of Western Australia, Perth, Western Australia 6009, Australia

⁵Department of Biology, Duke University, Durham, North Carolina 27708, USA

⁶Department of Genetics, University of Georgia, Athens, Georgia 30602, USA

⁷Howard Hughes Medical Institute, Duke University, Durham, North Carolina 27708, USA

⁸Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, CA 92037, USA

Abstract

DNA methylation is an epigenetic modification that differs between plant organs and tissues, but the extent of variation between cell types is not known. Here, we report single-base resolution whole genome DNA methylomes, mRNA transcriptomes, and small RNA transcriptomes for six cell populations covering the major cell types of the *Arabidopsis* root meristem. We identify widespread cell type specific patterns of DNA methylation, especially in the CHH sequence

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Corresponding authors: Ryan Lister, Philip N. Benfey, Joseph R. Ecker

⁹These authors contributed equally to this work.

Accession numbers

All sequence data can be downloaded from NCBI GEO under accession XXXXX for MethylC-seq data and XXXXX for RNA-seq data, and can be viewed at http://neomorph.salk.edu/Arabidopsis_root_methylomes.php.

Author contributions

P.N.B., J.R.E. and R.L. designed and supervised research. N.B., X.H. and M.V. collected cells. T.K., R.L., J.R.N. and M.A.U. conducted MethylC-seq experiments. R.L., J.R.N. and M.A.U. conducted RNA-seq experiments. T.K. and R.L. performed sequencing data processing. T.K., R.L., R.J.S. and T.S. performed statistical and bioinformatic analyses. R.J.S. performed imaging analyses. P.N.B., J.R.E., T.K., R.L. and T.S. prepared the manuscript.

Competing financial interests

The authors declare no competing financial interests.

context. The genome of the columella root cap is the most highly methylated Arabidopsis cell characterized to date. It is hypermethylated within transposable elements, accompanied by increased abundance of transcripts encoding RNA-directed DNA methylation (RdDM) pathway components and 24 nt small RNAs. Absence of the nucleosome remodeler DECREASED DNA METHYLATION 1, required for maintenance of DNA methylation, and low abundance of histone transcripts involved in heterochromatin formation suggests a loss of heterochromatin may occur in the columella, thus allowing access of RdDM factors to the whole genome, and producing excess 24 nt small RNAs in this tissue. Together, these maps provide new insights into the epigenomic diversity that exists between distinct plant somatic cell types.

Introduction

DNA methylation is an epigenetic modification of cytosine bases implicated in gene regulation. In plants, DNA methylation occurs in three distinct cytosine contexts; CG, CHG and CHH, where H is A, C or T. CG and CHG methylation is stably maintained by DNA METHYLTRANSFERASE 1 (MET1) and CHROMOMETHYLASE 3 (CMT3), respectively. *De novo* DNA methylation is catalyzed by DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) in all three sequence contexts, in a process that is guided by 24 nucleotide (nt) small RNA (smRNA), known as RNA-directed DNA methylation (RdDM)^{1,2}. DNA methylation may also be maintained independently of the RdDM pathway through the concerted action of DECREASED DNA METHYLATION 1 (DDM1) and CHROMOMETHYLASE 2 (CMT2)^{3,4}. DDM1 functions to displace the linker histone H1 in heterochromatic regions of the genome, allowing CMT2 access to the DNA, where it is able to catalyse the methylation of cytosines in the CHG and CHH contexts^{3,4}. While DNA methylation can be a stable epigenetic mark, faithfully maintained over many hundreds of generations⁵, dynamic changes in DNA methylation patterns can be observed over short time scales in response to the environment^{6,7}, or in different cell types of a single individual⁸⁻¹¹, presumably a result of differential regulation of the RdDM or CMT2-mediated DNA methylation pathways. Thus, DNA methylation is a stable but reversible epigenetic modification, and may reflect, or play an important role in maintaining, cell-type identity. However, further investigation is needed to characterize the epigenome in distinct cell types in order to investigate the potential role of any differences.

In Arabidopsis, a major biological role of DNA methylation is in silencing transposable element (TE) transcription. Loss of DNA methylation due to mutations in *DDM1* or *MET1* is sufficient for transcriptional activation of demethylated TE sequences, and transposition of some of these activated TEs^{3,12,13}. Although TE insertions may contribute to novel modes of gene regulation, excess TE activity produces deleterious mutations, and efficient TE silencing is crucial for the maintenance of genome integrity. Plants may be most vulnerable to TE activity in the stem cells, as these are the progenitor cells from which all others derive, and TE insertions within the stem cells will therefore be inherited by all descendant cells. Indeed, highly complex mechanisms of TE silencing have been reported in the sperm and embryo. TE silencing in the sperm is thought to be assisted by 21 nt smRNAs derived from the vegetative cell nucleus, a non-generative companion to the sperm, and in the developing embryo by endosperm-derived 24 nt smRNAs⁹⁻¹¹, indicating that silencing of TEs may be

particularly important in these cells. Plants have stem cell niches at distal axes, known as the shoot apical meristem (SAM) and root apical meristem (RAM). RdDM factors, DNA methyltransferases, and DDM1 are all upregulated to reinforce TE silencing in the SAM¹⁴. While there is some indication that gross levels of DNA methylation may be distinct in the RAM^{15,16}, patterns of DNA methylation in the RAM have not been studied at high resolution, and the dynamics of DNA methylation mediated TE silencing in the RAM are so far unexplored. Here we describe comprehensive DNA methylation and transcriptome profiling of six distinct cell types of the Arabidopsis RAM, revealing unique cell-type specific characteristics of DNA methylation and the machinery responsible for shaping the methylome.

Results

Columella is the most CHH hypermethylated cell type in Arabidopsis

To investigate patterns of DNA methylation in different plant cell types, we used protoplasting followed by fluorescence activated cell sorting (FACS) of cell populations marked by green fluorescent protein (GFP) in a range of reporter lines. These lines represent the major cell types or tissues in the root: epidermis (ProWER:GFP), cortex (ProCOR:GFP), endodermis (ProSCR:GFP), stele (ProWOL:GFP), whole columella root cap (PET111 enhancer trap line), and lower columella (ProCYCD5:GFP) (Fig. 1a). Two independently generated reporter lines were analyzed for the endodermis. Following isolation of highly enriched populations of each cell type (Fig. S1), we generated single-base resolution maps of cytosine methylation by whole genome bisulfite sequencing, and transcriptome profiles by RNA-seq and smRNA-seq (Fig. 1b and Table S1). Analysis of global levels of DNA methylation in the six cell populations revealed that methylation in all sequence contexts (mCG, mCHG, mCHH) were higher in the columella, with dramatically increased levels of mCHH (Fig. 1c). Comparison with previously published Arabidopsis methylomes showed that mCHH levels in the columella are higher than in any other tissue or cell type analyzed to date^{10,11} (Fig. 1c). The enrichment of mCHH in the columella was the most pronounced in the pericentromeric regions of the chromosome (Fig. 1d). Whole root tips from the PET111 transgenic line, as well as from Col-0, showed similar patterns and levels of mC as the non-columella cell types (Fig. 1c and d), indicating that the differences observed in the columella cell populations were due to cell type, and not a widespread perturbation of DNA methylation in the transgenic lines used for cell isolation.

Columella hypermethylation is the major source of widespread differential DNA methylation in the root meristem

To further investigate the large differences in DNA methylation patterns, we identified differentially methylated regions (DMRs) in the genome between the cell types. With a target false discovery rate of 5%, we identified 38,307 DMRs between the different cell types (Fig. 2a). Of these, 13.6% (5,225) were differentially methylated only in the CG context (CG-DMRs), while 82.9% (31,761) were differentially methylated only in the CH context (CH-DMRs) (Fig. 2a, Table S2 and S3). Regions differentially methylated in both the CG and CH context (C-DMRs) were rare, with only 1,321 such regions observed (Fig. 2a and Table S4). DMR length also seemed to be associated with DNA methylation context,

with CG-DMRs being, on average, shorter than CH- and C-DMRs (Fig. 2b). Overall, 13.8% of the nuclear genome was differentially methylated between the six cell types, mostly in the CH context (Fig. 2c).

Some regions of the genome are prone to spontaneous changes in DNA methylation levels^{17,18}. To determine if the regions of differential DNA methylation between cell types were due to spontaneous fluctuations in DNA methylation levels between the different transgenic lines used, we compared the root cell type specific DMRs with two types of previously identified spontaneous DMRs; transgenerational DMRs¹⁷ and population DMRs¹⁹. We found that 76% and 60% of root cell type specific CG-DMRs and C-DMRs, respectively, overlapped with population DMRs, whereas only 5% and 2% of root cell type specific CG- and C-DMRs overlapped with transgenerational CG- and C-DMRs (Fig. S2). We conclude that the majority of root cell type specific DMRs occur in regions of the genome known to be epigenetically labile, likely due to variation in smRNAs.

To determine if the enrichment of DNA methylation in pericentromeric regions (Fig. 1d) was linked to DMRs, we assessed the distribution of DMRs along the chromosomes (Fig. 2d and S3). While CG-DMRs are most abundant in the chromosome arms, the number of CH- and C-DMRs peaked in the proximal and distal pericentromeric regions, respectively. Closer inspection of the genomic features intersecting each set of DMRs revealed that over 80% of CG-DMRs overlapped with protein-coding gene bodies (Fig. 2e), while 73% of CH-DMRs and 44% of C-DMRs overlapped with TEs. The remaining CH-DMRs and C-DMRs were found to overlap mainly with intergenic regions or pseudogenes.

Hierarchical clustering based on differences in DNA methylation showed that the columella cells form a highly distinct group compared to other cells of the root (Fig. 2f). Interestingly, DNA methylation patterns seemed to be more similar between cell types located physically close to one another in the root, regardless of their lineage, whereas transcriptional profiles were more dependent on cell lineage than physical position in the root (Fig. S4). This may suggest that methylation patterns are in part regulated by positional information or cell-cell communication. Columella cells were highly distinct in their DNA methylation landscape, particularly in the mCHH context. Methylation at CH- and C-DMRs was higher in the columella than in other cell types, suggesting that CHH hypermethylation in the columella is the primary basis for CH- and C-DMRs among root meristem cells (Fig. 2g and S5).

As mCHH is deposited by two distinct DNA methyltransferases, DRM2 and CMT2^{3,4}, we sought to determine which methyltransferase was responsible for mediating changes in mCHH in each set of DMRs. We analysed mCHH levels within DMR coordinates in leaves of wild type, *drm1drm2*, and *cmt2* plants in order to categorize DMRs as DRM2 or CMT2 targets, using previously published DNA methylation data²⁰ (Fig. S6). For CH-DMRs, both *drm1drm2* and *cmt2* showed decreased mCHH in these regions, but the effect of *cmt2* was much larger, whereas for C-DMRs only *drm1drm2* caused a decrease in mCHH levels. These results reveal that mCHH within CH-DMRs and C-DMRs is mainly catalyzed by CMT2 and DRM2, respectively. DRM2 is involved in two types of RdDM, the canonical Pol IV-mediated RdDM guided by 24 nt smRNAs^{1,2}, and RDR6-mediated RdDM guided by 21 and 22 nt smRNAs²¹. We detected upregulation of 21–24 nt smRNA abundance within both

CH-DMRs and especially C-DMRs in the columella, but 24 nt smRNA were predominant (Fig. 2h and S7), suggesting that the canonical Pol IV-mediated RdDM pathway plays a major role in establishing these DMRs. We did not observe higher steady state transcript abundance of TEs in the columella (Fig. S8).

Gene body methylation in the CG context is correlated with constitutive gene expression^{22–24}. In contrast, DNA methylation in gene-flanking regions is thought to repress gene expression. To address whether DMRs affect the expression of nearby genes, we correlated DMR methylation levels and nearby gene expression levels (Fig. 2i). Most CG-DMRs were located within the gene bodies, especially near transcriptional termination sites. However, minimal correlation between methylation levels at CG-DMRs and expression levels of nearby genes was observed. CH- and C-DMRs are largely excluded from the gene bodies. While the correlation between CH- and C-DMR methylation and gene expression was also variable, methylation at transcription start sites was weakly negatively correlated with the transcript abundance of nearby genes. Similarly, methylation at C-DMRs within gene bodies showed a negative correlation with gene expression. These results suggest root cell type specific CH and C-DMRs are only weakly associated with cell type specific gene expression patterns. Additionally, gene ontology enrichment analysis showed that CH-DMR associated genes were enriched for response genes, such as “defense response” and “innate immune response” (Fig. S9). This suggests that CH-DMRs only weakly correlate with nearby gene expression, and may only have an impact on gene expression under specific environmental circumstances.

Transposable elements are targets for CHH hypermethylation

Although only a small percentage of CH-DMRs were found to intersect with gene bodies (Fig. 2f), these still represented over 1,000 gene loci due to the abundant nature of CH-DMRs. To further investigate whether there was a correlation between mCHH levels within genes and the transcript abundance of those genes, we ordered all TAIR10 genes based on the average transcript abundance among cell populations and further analyzed patterns of DNA methylation (Fig. 3). This revealed that, while levels and patterns of mCG and mCHG were similar between cell types (Fig. 3a), lowly expressed and silent genes were CHH hypermethylated in the columella. Furthermore, we found that the number of genes harboring TEs were also enriched in genes with lower expression (Fig 3a), suggesting that increases in mCHH within lowly expressed genes may be due to the hypermethylation of TEs contained within these genes. As mCHH serves to transcriptionally silence TEs in Arabidopsis, and most CH-DMRs were found within annotated TEs, we compared patterns and levels of DNA methylation across all TEs in the genome (Fig. 3b–d). Levels of mCG and mCHG in TEs were only moderately higher in both of the columella cell populations, consistent with our observations on a genome-wide scale (Fig. 1c). However, a large increase in mCHH in TEs in both of the columella cell populations was observed compared with the other cell types, and this was consistent across all known TE superfamilies in Arabidopsis (Fig. S10). This indicates that, while some CH-DMRs were found to intersect with protein-coding genes, differences in mCHH between cell types can be attributed almost entirely to the CHH hypermethylation of TEs in the columella. As TEs

are greatly enriched in the pericentromeric heterochromatin, this would also explain the enrichment of mCHH and CH-DMRs in the pericentromeric regions (Fig. 1c and 2d).

Enhanced RNA-directed DNA methylation in the columella

As we observed an increase in mCHH in TEs, as well as an increase in 24 nt smRNA abundance at CH-DMRs, we next sought to determine whether there might be transcriptional upregulation of the RdDM pathway in the columella. Analysis of the RNA-seq data revealed an increase in transcripts encoding components of the RdDM pathway in the columella as compared to the other cell populations (Fig. 4a). In particular, we found an enrichment for transcripts encoding proteins needed for smRNA biogenesis, such as the major unique Pol IV component *NRPD1a*, as well as *CLSY1*, *RDR2* and *DCL3*, while those components involved directly in the deposition of DNA methylation were only mildly upregulated in the columella^{25–28}. To investigate whether this increased production of smRNA biogenesis machinery in the columella translated into an increase in the proportion of 24 nt smRNAs sequenced, we assessed levels of uniquely mapped 21, 22, 23, and 24 nt smRNAs genome-wide (Fig. 4b and Table S5). This revealed a strong increase in the fraction of 24 nt smRNAs in the columella, indicating that the hypermethylation of TEs in the columella is coupled with the transcriptional upregulation of the smRNA biogenesis machinery and increased production of 24 nt smRNAs needed for the RdDM pathway.

DDM1 protein is not present in the columella

In the vegetative cell nucleus of the pollen, a loss of mCG and mCHH throughout the genome is coupled with increased mCHH at the centromere, the absence of DDM1 protein, and loss of heterochromatin^{9,10} (Fig. S11). This triggers TE transcriptional activation and increased production of 21 nt smRNAs from TE transcripts, which are thought to be transported to the sperm cells to reinforce TE silencing in the germline^{9,10}. We observed an increase in 24 nt smRNA and CHH hypermethylation of TEs in columella cells. Although no decrease in *DDM1* transcript abundance specific to the columella was detected (Fig. 4a and 5a), analysis of a transgenic line expressing the DDM1-GFP fusion protein revealed that DDM1-GFP was undetectable in the columella, whereas it was present in the nuclei of other root cell types (Fig. 5b). This indicates that *DDM1* is transcribed in the columella, but either the transcripts are not translated or there is rapid degradation of DDM1 protein. Despite an apparent lack of DDM1 in the columella, and in contrast to *ddm1*, normal levels of mCG and mCHG are maintained at TEs, and there are elevated levels of mCHH (Fig. 3d, 5c).

DDM1-dependent mCHH deposition is catalyzed by the DNA methyltransferase CMT2³, and the RdDM pathway together with CMT2 are responsible for almost all mCHH in the genome⁴. We classified all methylated TEs into four clusters, based on the mCHH levels within TE bodies of wild type, *drm1 drm2*, *cmt2* and *ddm1* leaf tissue (Fig. 5c left panel). mCHH levels in TEs in clusters 1 and 2 were decreased in *drm1 drm2*, indicating that they were RdDM-dependent. mCHH levels in TEs in cluster 3 and cluster 4 were decreased in *cmt2*, indicating that their methylation was CMT2-dependent. Strikingly, in the columella, TEs in all four clusters were hypermethylated (Fig. 5c). RdDM-dependent TEs were hypermethylated, accompanied by 24 nt, but not 21 nt, smRNA accumulation. CMT2-dependent TEs were hypermethylated in the columella, and those located in chromosome

arms were accompanied by 24 nt smRNA accumulation, consistent with 24 nt smRNA enrichment in CH-DMRs. The edges of CMT2-dependent TEs are subjected to RdDM, and 24 nt smRNAs are enriched in these regions³. However, the edges as well as the bodies of CMT2-dependent TEs accumulated 24 nt smRNA in the columella (Fig. 5c and d), suggesting that the bodies of CMT2-dependent TEs are also subjected to RdDM. This may account for CHH hypermethylation of CMT2-dependent TEs, but lower expression of *CMT2* in the columella (Fig. 4a).

DDM1 is normally required for the displacement of histone H1 at heterochromatic regions of the genome, allowing DNA methyltransferases MET1, CMT3 and CMT2 to access and methylate the DNA³. As loss of H1 suppresses the reduction in DNA methylation in *ddm1* mutants, we examined transcript levels for the two canonical histone H1 genes, *H1.1* and *H1.2*, and observed lower abundance of transcripts for both genes in columella cells as compared with other cells in the root meristem (Fig. 4a). Also *H2A.W6* and *H2A.W7*, which are required for chromatin condensation²⁹, were down-regulated in the columella (Fig. 4a), suggesting that the columella may lose heterochromatin through a reduction of heterochromatin related components. Loss of heterochromatin in the columella may play a role in enhancing generation of the 24 nt smRNA transcripts needed for RdDM, leading to the observed CHH hypermethylation of TEs.

Discussion

Plants are complex multicellular organisms that contain a broad variety of cell types with specialized functions. While differences in patterns of DNA methylation have been observed previously between different somatic tissues¹⁶ and reproductive cell types^{9-11,30}, this is the first report of differences in DNA methylation between cell types from the same somatic tissue. The root meristem contains a diverse variety of cell types. Among these, the columella is a group of specialized gravity-sensing cells, required for proper development of the root^{31,32}. The most striking and unique feature of these root cell methylomes is the CHH hypermethylation of TEs in the columella, which is obscured when methylomes are analyzed using whole roots or dissected root tips. The majority of detected DMRs are columella hypermethylated CH-DMRs occurring in epigenetically labile regions. To date, the columella is the most highly methylated tissue or cell type characterized in Arabidopsis, and such an extreme level of hypermethylation is not recapitulated by any known gene silencing mutant. Columella cells are rapidly replaced in the root as they grow outward from the columella initials, located below the quiescent centre (QC), and are ultimately detached into the soil. Consequently, columella cells are short lived and undergo rapid differentiation. mCHH primarily serves to silence TE transcription in Arabidopsis, preventing potentially damaging genetic mutations caused by transposition. Due to the terminally differentiated and short-lived nature of columella cells, it is reasonable to expect that TE insertions in the columella would have little impact on root function and fitness, as these mutations would be quickly lost from the plant. The apparent enhancement of DNA methylation-mediated TE silencing in the columella is therefore counterintuitive. One possible explanation for this seeming contradiction is that the columella acts as a companion to nearby stem cells, and is similar in function to the reproductive companion cells found in the developing pollen and seed^{9-11,30}. Excess 24 nt smRNAs produced by the columella, required for initiation of

RdDM at TEs, may be transported into the neighboring stem cell niches such as the QC, reinforcing transcriptional silencing of TEs in these stem cells in a manner analogous to the smRNA transport thought to occur between the vegetative cell nucleus and sperm cells in the developing pollen, between the central cell and egg, and possibly between the endosperm and embryo^{10,11}. As the root stem cells are responsible for the establishment of tissue patterning in the root, and all cells of the root descend from these stem cells, transpositions in them may have a larger impact on plant fitness, and therefore have a greater need for effective TE silencing, similar to generative stem cells in the germline. Although DDM1 is undetectable in both the columella and the vegetative cell nucleus of the pollen, the patterns of mCHH at TEs are distinct between these two cell types, with mCHH levels in the vegetative cell nucleus being more similar to that found in *ddm1*. Heterochromatin has been suggested to inhibit RdDM, while open chromatin increases the accessibility of RdDM components to the genome leading to hypermethylation³³. DDM1, and associated proteins such as H1, may play an important role in regulating the exclusion of RdDM-related factors from the heterochromatin, and it is possible that DDM1 protein accumulation is actively suppressed in the columella to allow RdDM at DDM1-regulated heterochromatic regions. Another possible explanation for the CHH hypermethylation and upregulation of the RdDM pathway in the columella may be that special attributes of the columella, such as the rapid differentiation after one division of the stem cell and possible increased ploidy level³⁴, are conducive to hypermethylation of TEs. Future experiments will be needed to further examine these hypotheses.

Materials and methods

Cell isolation

Seedlings were grown vertically for 6 days after plating on 1× Murashige and Skoog media supplemented with 1% sucrose and 1% agar. All seedlings were grown under standard long day conditions (16 hours of light, 8 hrs of darkness, 22 °C). Fluorescent Activated Cell Sorting (FACS) was performed using cell specific GFP lines as described previously³⁵. The columella root cap was marked with the enhancer trap PET111³⁶, the bottom two layers of the columella were marked with ProCYCD5:GFP³⁷, the stele with ProWOL:GFP³⁸, the endodermis with ProSCR:GFP³⁹, the cortex with ProCORTEX:GFP⁴⁰, and both the epidermis and lateral root cap with ProWER:GFP⁴¹. Sorted cells were collected directly into specific lysis buffers that were compatible with downstream applications. Cells used for bisulfite sequencing, mRNA-seq, smRNA-seq were lysed in Buffer AP1 (Qiagen), Buffer RLT (Qiagen), Trizol (Invitrogen). All samples were immediately stored at −80 °C until gDNA and RNA was extracted using DNeasy Plant mini kit (Qiagen) and RNeasy Plant mini kit (Qiagen) or Trizol, respectively.

MethylC-seq

MethylC-seq library preparation, read mapping, base calling were performed as described previously^{42–44}, except that reads were mapped against C-to-T converted TAIR10 reference genome and library amplification was performed with either KAPA HiFi U+ (KAPA) or PfuTurboCx enzyme (Agilent). Bisulfite non-conversion rate was estimated from the total

number of cytosine base calls divided by the total coverage at cytosine positions in the naturally unmethylated chloroplast genome.

Identification of differentially methylated regions

Differentially methylated regions (DMRs) were identified using the methylpy pipeline⁴⁵. Briefly, differentially methylated sites (DMSs) were identified by root mean square tests with false discovery rate at 0.05, using 1,000 permutations. Cytosine positions at least with 4 reads were examined for differential methylation. Then, DMSs within 200 bp were collapsed into DMRs. DMRs were classified into CG-DMRs (only CG difference), CH-DMRs (only CHG and/or CHH difference), C-DMRs (CG and CHG and/or CHH difference). In addition, CG-DMRs, CH-DMRs, C-DMRs with fewer than 5, 5, 10 DMSs, respectively, were discarded in following analysis. Differential methylation tests were performed among samples, not in pairwise manner, generating a set of all non-redundant DMRs among the samples. Methylation levels of each region were calculated as weighted methylation levels⁴⁶, in which methylation level was equal to the frequency of C base calls at C positions within the region divided by the frequency of C and T base calls at C positions within the region.

RNA-seq

RNA-seq library preparation was performed using the Illumina TruSeq RNA Library Prep kit from polyA+ selected mRNA as per manufacturer's instructions. smRNA sequencing data was obtained from a previous study⁴⁷. smRNA data were processed and mapped to the TAIR10 genome as described previously⁴⁸. smRNAs levels were normalized to TE size and library size by counting reads per kilobase of TE per million reads mapped (RPKM). Only reads that mapped uniquely to the genome contributed to the average count for each TE. RNA-seq data were mapped to the TAIR10 reference genome using Tophat2 with the default parameters⁴⁹, and quantified using Cuffdiff⁵⁰.

Associating DMRs with proximal genes

DMRs located within 3 kb of gene upstream regions, gene bodies, and 3 kb of gene downstream regions were extracted, and relative position to genes were assigned by the middle position of DMRs. Some DMRs were located within multiple genomic features, for example in the 3kb upstream regions, gene bodies, or 3kb downstream regions for more than one gene. We refer to all possible pairwise comparisons between DMRs and nearby genomic features as "combinations". Pearson correlation coefficients between the methylation levels of DMRs and the expression levels of proximal genes (FPKM) were computed and plotted as density.

Clustering TEs

mCHH levels within annotated TE bodies at least 400 bp in length were computed, and only TEs with at least 10% mCHH at least in one sample from Col-0, *drm1 drm2*, *cmt2*, and *ddm1* were assigned as methylated TEs. TEs were then clustered into four clusters by using R *k*-means function, with the "centers" parameter set to 4.

Microscopy analysis

The DDM1-GFP transgenic line was described previously⁹. Seeds were plated on 1/2× Linsmaier and Skoog media. Seedlings three days after germination were incubated in propidium iodide for 5 min to stain cell walls of root tips, and imaged using Zeiss LSM 710 Confocal Microscope.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Keith Slotkin and James A. H. Murray for kindly providing DDM1-GFP seeds and ProCYCD5:GFP seeds, respectively. T.K. was supported by the Japan Society for the Promotion of Sciences Research Abroad Fellowship. T.S. was supported by the Jean Rogerson Postgraduate Scholarship. This research was supported by grants from the National Science Foundation (MCB-1344299 to J.R.E and IOS-1021619 to P.N.B.), by the National Institutes of Health (GM R01-043778 to P.N.B.) and by the Gordon and Betty Moore Foundation (GBMF3034 to J.R.E and GBMF3405 to P.N.B.). R.L. was supported by the Australian Research Council (FT120100862). R.J.S. was supported by the National Institutes of Health (R00GM100000). J.R.E and P.N.B. are investigators of the Howard Hughes Medical Institute.

References

1. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature reviews. Genetics*. 2010; 11:204–220.
2. Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature reviews. Genetics*. 2014; 15:394–408.
3. Zemach A, et al. The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*. 2013; 153:193–205. [PubMed: 23540698]
4. Stroud H, et al. Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nature structural & molecular biology*. 2014; 21:64–72.
5. Cubas P, Vincent C, Coen E. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature*. 1999; 401:157–161. [PubMed: 10490023]
6. Secco D, et al. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. *eLife*. 2015; 4
7. Downen RH, et al. Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:E2183–E2191. [PubMed: 22733782]
8. Pignatta D, et al. Natural epigenetic polymorphisms lead to intraspecific variation in Arabidopsis gene imprinting. *eLife*. 2014; 3:e03198. [PubMed: 24994762]
9. Slotkin RK, et al. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell*. 2009; 136:461–472. [PubMed: 19203581]
10. Calarco JP, et al. Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell*. 2012; 151:194–205. [PubMed: 23000270]
11. Ibarra CA, et al. Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science*. 2012; 337:1360–1364. [PubMed: 22984074]
12. Mirouze M, et al. Selective epigenetic control of retrotransposition in Arabidopsis. *Nature*. 2009; 461:427–430. [PubMed: 19734882]
13. Tsukahara S, et al. Bursts of retrotransposition reproduced in Arabidopsis. *Nature*. 2009; 461:423–426. [PubMed: 19734880]
14. Baubec T, Finke A, Mittelsten Scheid O, Pecinka A. Meristem-specific expression of epigenetic regulators safeguards transposon silencing in Arabidopsis. *EMBO reports*. 2014; 15:446–452. [PubMed: 24562611]

15. Seymour DK, Koenig D, Hagmann J, Becker C, Weigel D. Evolution of DNA methylation patterns in the Brassicaceae is driven by differences in genome organization. *PLoS genetics*. 2014; 10:e1004785. [PubMed: 25393550]
16. Widman N, Feng S, Jacobsen SE, Pellegrini M. Epigenetic differences between shoots and roots in *Arabidopsis* reveals tissue-specific regulation. *Epigenetics*. 2014; 9:236–242. [PubMed: 24169618]
17. Schmitz RJ, et al. Transgenerational epigenetic instability is a source of novel methylation variants. *Science*. 2011; 334:369–373. [PubMed: 21921155]
18. Becker C, et al. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature*. 2011; 480:245–249. [PubMed: 22057020]
19. Schmitz RJ, et al. Patterns of population epigenomic diversity. *Nature*. 2013; 495:193–198. [PubMed: 23467092]
20. Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell*. 2013; 152:352–364. [PubMed: 23313553]
21. Nuthikattu S, et al. The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21–22 nucleotide small interfering RNAs. *Plant physiology*. 2013; 162:116–131. [PubMed: 23542151]
22. Zhang X, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell*. 2006; 126:1189–1201. [PubMed: 16949657]
23. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nature genetics*. 2007; 39:61–69. [PubMed: 17128275]
24. Tran RK, et al. DNA methylation profiling identifies CG methylation clusters in *Arabidopsis* genes. *Current biology : CB*. 2005; 15:154–159. [PubMed: 15668172]
25. Xie Z, et al. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol*. 2004; 2:E104. [PubMed: 15024409]
26. Onodera Y, et al. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell*. 2005; 120:613–622. [PubMed: 15766525]
27. Henderson IR, et al. Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. *Nature genetics*. 2006; 38:721–725. [PubMed: 16699516]
28. Smith LM, et al. An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. *The Plant cell*. 2007; 19:1507–1521. [PubMed: 17526749]
29. Yelagandula R, et al. The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in *Arabidopsis*. *Cell*. 2014; 158:98–109. [PubMed: 24995981]
30. Hsieh TF, et al. Genome-wide demethylation of *Arabidopsis* endosperm. *Science*. 2009; 324:1451–1454. [PubMed: 19520962]
31. Tsugeki R, Fedoroff NV. Genetic ablation of root cap cells in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*. 1999; 96:12941–12946. [PubMed: 10536027]
32. Swarup R, et al. Root gravitropism requires lateral root cap and epidermal cells for transport and response to a mobile auxin signal. *Nature cell biology*. 2005; 7:1057–1065. [PubMed: 16244669]
33. Schoft VK, et al. Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. *EMBO reports*. 2009; 10:1015–1021. [PubMed: 19680290]
34. Dolan L, et al. Cellular organisation of the *Arabidopsis thaliana* root. *Development*. 1993; 119:71–84. [PubMed: 8275865]
35. Iyer-Pascuzzi AS, Benfey PN. Fluorescence-activated cell sorting in plant developmental biology. *Methods Mol Biol*. 2010; 655:313–319. [PubMed: 20734270]
36. Nawy T, et al. Transcriptional profile of the *Arabidopsis* root quiescent center. *The Plant cell*. 2005; 17:1908–1925. [PubMed: 15937229]

37. Collins C, Dewitte W, Murray JA. D-type cyclins control cell division and developmental rate during *Arabidopsis* seed development. *Journal of experimental botany*. 2012; 63:3571–3586. [PubMed: 22412186]
38. Mahonen AP, et al. A novel two-component hybrid molecule regulates vascular morphogenesis of the *Arabidopsis* root. *Genes & development*. 2000; 14:2938–2943. [PubMed: 11114883]
39. Birnbaum K, et al. A gene expression map of the *Arabidopsis* root. *Science*. 2003; 302:1956–1960. [PubMed: 14671301]
40. Lee JY, et al. Transcriptional and posttranscriptional regulation of transcription factor expression in *Arabidopsis* roots. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103:6055–6060. [PubMed: 16581911]
41. Lee MM, Schiefelbein J. WEREWOLF, a MYB-related protein in *Arabidopsis*, is a position-dependent regulator of epidermal cell patterning. *Cell*. 1999; 99:473–483. [PubMed: 10589676]
42. Lister R, et al. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell*. 2008; 133:523–536. [PubMed: 18423832]
43. Lister R, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*. 2011; 471:68–73. [PubMed: 21289626]
44. Urich MA, Nery JR, Lister R, Schmitz RJ, Ecker JR. MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. *Nature protocols*. 2015; 10:475–483. [PubMed: 25692984]
45. Schultz MD, et al. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature*. 2015; 523:212–216. [PubMed: 26030523]
46. Schultz MD, Schmitz RJ, Ecker JR. 'Leveling' the playing field for analyses of single-base resolution DNA methylomes. *Trends in genetics : TIG*. 2012; 28:583–585. [PubMed: 23131467]
47. Breakfield NW, et al. High-resolution experimental and computational profiling of tissue-specific known and novel miRNAs in *Arabidopsis*. *Genome research*. 2012; 22:163–176. [PubMed: 21940835]
48. Lister R, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009; 462:315–322. [PubMed: 19829295]
49. Kim D, et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology*. 2013; 14:R36. [PubMed: 23618408]
50. Trapnell C, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols*. 2012; 7:562–578. [PubMed: 22383036]

Figure 1. Cell type specific patterns of DNA methylation in the root meristem

(a) Schematic representation of the six root cell types used in this study. Endodermis has two independent replicates (indicated by numerals) for MethylC-seq and RNA-seq. *: MethylC-seq data only. (b) A genome browser snapshot showing DNA methylation level, RNA-seq reads, smRNA-seq reads. (c) Global levels of DNA methylation in each context for root and reproductive cells (VN¹¹: vegetative nucleus, SP¹¹: sperm, MS¹⁰: microspore, EM³⁰: embryo, EN³⁰: endosperm). (d) Heatmap showing mC levels within 100 kb bins and genes and TEs within 50 kb bins over the entire genome. Maximum mC levels are 0.91 (mCG), 0.72 (mCHG), 0.34 (mCHH).

Figure 2. Differentially methylated regions among six root cell types

(a) Numbers of each type of DMR. (b) Average sizes of each type of DMR. (c) Genomic fraction of each type of DMR relative to the whole genome. (d) Genome wide distribution of each type of DMR. Counts were scaled by maximum count as one. (e) Genomic features covering DMRs. (f) Hierarchical clustering of six root cell types for CG-, CH-, and C-DMRs. (g) mCHH levels within CH- and C-DMRs. (h) 24 nt smRNA expression levels in CH- and C-DMRs. (i) Correlation between DMR methylation and gene expression levels. “Combinations” refers to the number of all possible comparisons between DMRs and the nearby genomic features. Please see methods for details.

Figure 3. DNA methylation in genes and TEs

(a) DNA methylation patterns within genes ordered by average mRNA abundance. (b) DNA methylation patterns and levels within TEs for each cell type. TEs show greatly increased mCHH in the columella genome, while mCG and mCHG levels are similar and moderately higher than other cell types.

Figure 4. Transcript levels of DNA methylation related genes
(a) Increased transcript abundance for genes involved in smRNA biogenesis in the columella. Scale is log2 fold change FPKM for each cell type compared to average FPKM for all cell types. Average FPKM for all cell types are also shown. (b) Fraction of uniquely-mapped smRNA for each size class relative to total uniquely-mapped smRNA. The proportion of 24 nt smRNAs relative to other size classes is greatly increased in the columella.

Figure 5. Loss of DDM1 in the columella

(a) DDM1 transcript abundance in all cell types. (b) Absence of DDM1-GFP in the columella. (c) mCHH levels and smRNA accumulation around methylated TEs. Methylated TEs were classified into 4 clusters based on TE body methylation levels of wild type, *drm1drm2*, *cmt2*, *ddm1* in leaf, by using *k*-means method (centres = 4). Left panel shows mCHH levels within TEs and their 2kb upstream and downstream regions. Each region consists of 40 equally sized bins. Right panel shows smRNA expression levels as in left

panel. TEs were ordered by coordinate within each cluster. (d) Representative genome browser snapshots for TEs in each cluster.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript