UCLA
UCLA Previously Published Works

Title

Epigenetic differences between shoots and roots in Arabidopsis reveals tissue-specific regulation

Permalink

https://escholarship.org/uc/item/0pz9k4d6

Journal

Epigenetics, 9(2)

ISSN

1559-2294

Authors

Widman, Nicolas
Feng, Suhua
Jacobsen, Steven E
et al.

Publication Date

2014-02-01

DOI

10.4161/epi.26869

Peer reviewed

Epigenetic differences between shoots and roots in *Arabidopsis* reveals tissue-specific regulation

Nicolas Widman, Suhua Feng, Steven E Jacobsen, and Matteo Pellegrini*

Department of Molecular, Cell and Developmental Biology; University of California, Los Angeles; Los Angeles, CA USA

DNA methylation and nucleosome densities play a critical role in the regulation of gene expression. While much is known about the mechanisms of transcriptional control that are mediated by these, less is known about the degree to which they are tissue-specific. By comparing DNA methylation, nucleosome densities and transcriptional levels in different tissue types we can gain a clearer understanding of the extent to which these mechanisms influence gene expression in a tissue specific manner. We compared DNA methylation in Arabidopsis shoots and roots and found extensive differences across the genome. We computed DNA methylation differences between roots and shoots at single cytosines and found that one in every 173 cytosines was differentially methylated. In addition we compared DNA methylation with tissue specific gene expression and nucleosome density measurements to identify associations between these. We also identified a group of genes that are strongly correlated with these epigenetic marks and are significantly differentially methylated between roots and shoots. These root-specific genes are part of the extensin family, and are preferentially methylated and have at least 10-fold higher expression and lower nucleosome density in roots relative to shoots.

Introduction

DNA methylation is known to have a regulatory effect on gene expression and chromatin structure.[1] It is one of several epigenetic mechanisms that include chromatin structure, histone protein modifications and RNA interference.[2] In eukaryotic DNA cytosine is the only nucleotide base that is methylated and its methylation is influenced by the sequence of the two downstream bases.[3] The most highly methylated cytosines are those that are immediately followed by guanine, referred to as CG sites. CG methylation is found in both genes and repeats and can be involved in the regulation of gene expression.[4] In plant genomes, there are 2 additional sequence contexts where methylation can occur: CHG and CHH (H represents any base other than guanine). Methylation at CHG and CHH sequence contexts, in contrast to CG methylation, is absent over genes and mostly found in intergenic, repeat-rich regions of the genome and plays a critical role in silencing transposons.[5] In addition to DNA methylation, changes in chromatin structure can also affect gene expression. These changes include modifications to histone proteins and the positioning of nucleosomes.

Despite extensive research in animals, which has described changes in DNA methylation during development,[6] the degree to which these epigenetic mechanisms determine tissue differentiation in plants is still poorly understood. Prior work has shown that global methylation differences between tissue types in rice are small, with CG methylation levels being nearly identical and non-CG methylation increasing slightly with the age of the tissue.[7] The greatest methylation differences found across rice tissue types were identified in the endosperm and consist of global non-CG hypomethylation along with hypomethylation of specific genes. Similarly, Arabidopsis endosperm and pollen genomes are significantly demethylated[5]

Nucleosome positioning has also been shown to have a significant effect on transcriptional regulation. Nucleosomes appear to play a role in determining the boundaries between exons and introns and are more strongly positioned over exons than introns regardless of the gene expression level.[8] Additionally, heterochromatin in pericentromeric regions is associated with different nucleosome marks than euchromatin. However, as in the case of DNA methylation, differences in nucleosome occupancies between plant tissues have not been well characterized.

To further extend our understanding of tissue-specific epigenetic differences in plants, and their functional roles, we measured DNA methylation, nucleosome positioning and gene expression in both Arabidopsis root and shoot tissues. Our goal was to identify tissue-specific DNA methylation and nucleosome distributions and associate these with changes in gene expression between differentiated tissue types. We sought to identify genes that may be regulated by DNA methylation during tissue differentiation resulting in tissue specific expression.
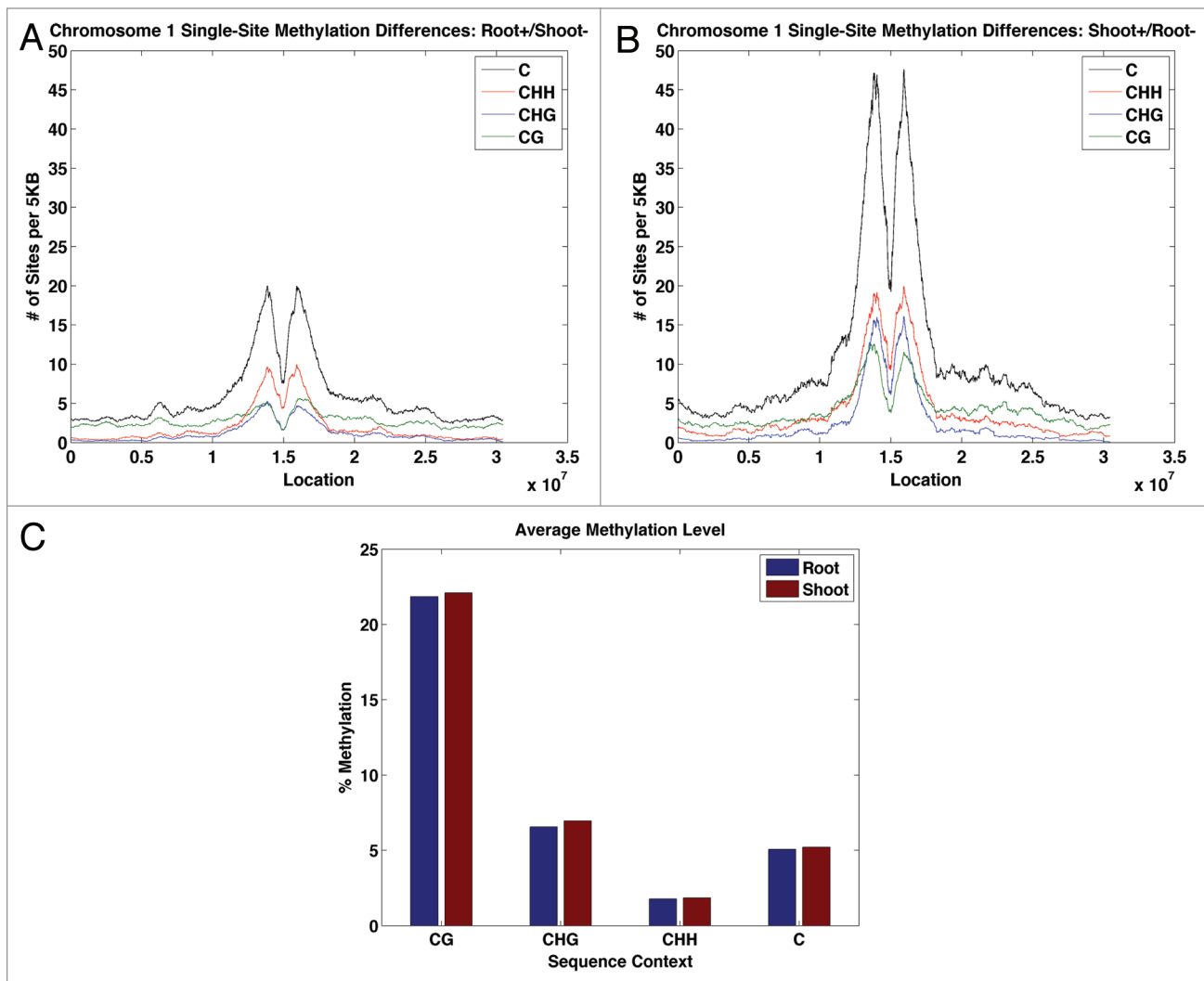
Figure 1. Genome-wide methylation levels. (**A**) Rate of single-site methylation differences by sequence context with higher methylation in root in chromosome 1. (**B**) Rate of single-site methylation differences by sequence context with higher methylation in shoot in chromosome 1. (**A and B**) X, position on chromosome; Y, number of instances of differential methylation per 5000 bases. (**C**) Average genome-wide methylation levels roots and shoots.

Results

DNA methylation in shoots and roots

We measured DNA methylation in both root and shoot samples using whole genome bisulfite sequencing. The shoot samples were the same as those we previously published,[4] except that we employed a slightly different sequencing library generation method (see Methods).[9] The root tissues were from root culture grown in liquid medium (see Methods). This root culture method was chosen because it provides a comprehensive representation of all the different types of cells from root and produces consistent results across replicates as we can sensitively control the environment under this growth condition.

Our reads were mapped to the Arabidopsis genomes using the method described by Seeker;[10] we obtained methylation levels by computing the fraction of converted cytosines. We were able to measure methylation levels for 75–80% of all cytosines. The methylation levels for the remaining 20% of cytosines could not be accurately measured due to inadequate coverage or certain regions of the genome not being mapped uniquely using our read length. The data may be visualized at our genome browser http://genomes.mcdb.ucla.edu and downloaded from GEO usingthe accession number GSE52762.

We analyzed the root and shoot methylomes by comparing the methylation state of each individual site in the 2 samples. We used a binomial model to identify sites that were differentially methylated (see Methods). Using this single-site differential methylation analysis, we found that 1.55% of sites passed our threshold. We also found that sites that are differentially methylated are preferentially hypermethylated in shoots compared with roots with 1.85 times as many hypermethylated sites in shoots than in roots. This effect is more significant in pericentromeric regions and at sites of non-CG methylation (**Fig. 1A and B**). In contrast, single-site methylation differences in non-centromeric regions are dominated by CG methylation. This is not surprising given that CG methylation is the
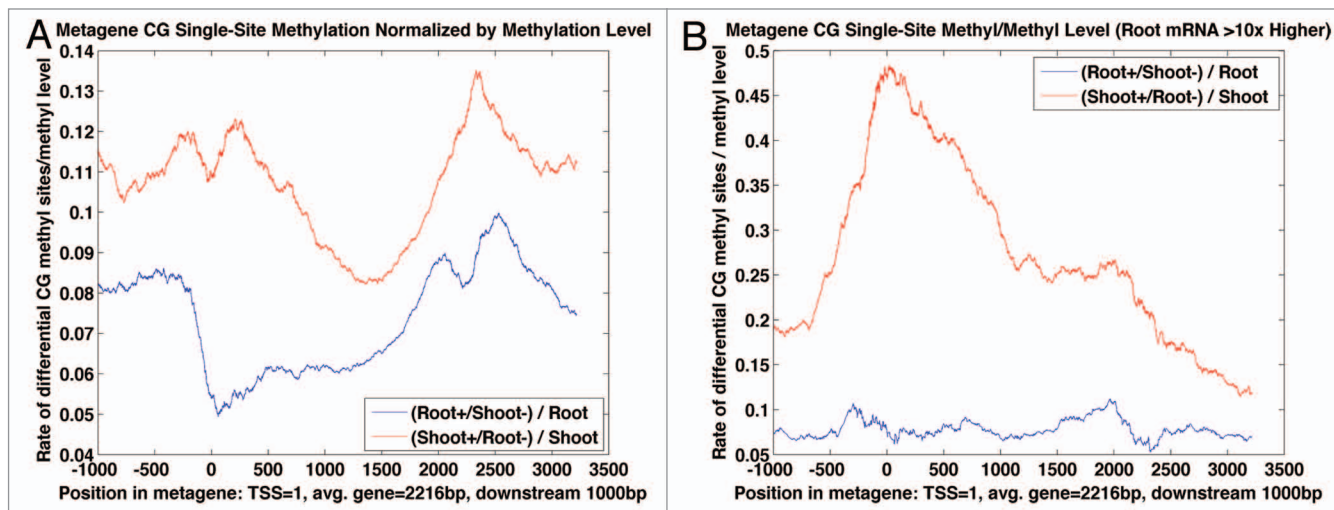
Figure 2. Metagene methylation differences. (**A**) Average rate of single-site methylation differences (CG sequence context only) across all genes. (**B**) Average rate of single-site methylation differences (CG sequence context only) across genes with 10× greater expression in roots. (**A and B**) X, metagene data normalized to average gene length of 2216 bases and 1000 bases upstream and downstream of gene, position on gene; Y, rate of differentially methylated sites divided by methylation level.

dominant context in euchromatic regions, while both CG and non-CG methylation is present in heterochromatic regions. We note, however, that global methylation levels do not differ significantly between roots and shoots with CG methylation being only 1 percent higher in shoots relative to roots and non-CG methylation being 5 percent higher in shoots relative to roots (**Fig. 1C**).

In previous studies, it was found that DNA methylation within genes is dominated by CG methylation.[4] Transcription start and termination sites are generally demethylated, as these regions are bound by regulatory factors that likely interfere with the ability of DNA methyltransferases to access and methylate DNA. CHG and CHH methylation levels generally show the opposite behavior as they are only found at very low levels within genes and at higher levels in intergenic, repeat-rich regions. We asked whether the density of single site methylation differences between shoots and roots varied across genes. To account for the changes in the average levels of methylation across genes, we normalized the densities by these averages. This normalization was performed to remove the expected bias wherein changes in highly methylated sites are more easily detected than changes to lowly methylated sites. We find that, within genes, differentially methylated sites are found primarily at CG sites and are preferentially methylated in shoot relative to root. Shoot hypermethylated sites occur about 1.85 times more frequently than sites preferentially methylated in root (**Fig. 2A**). Furthermore, we observe a higher fraction of these sites at the transcription initiation and termination boundaries, suggesting that these regions have more variable methylation levels than gene bodies.

Differential expression and DNA methylation

We obtained RNA-seq data from Arabidopsis shoots and roots using oligo-DT priming.[11] The reads were mapped to the genome using bowtie and the counts per gene were compared across the two samples. Overall, we identified 2424 genes that were significantly differentially expressed between the two tissues using the criteria that their fold change was at least 10.[11] Out of these differentially expressed genes, 1292 were more expressed in roots. These genes were enriched for functional groups associated with cell wall organization and biogenesis.

The root overexpressed transcripts were also strongly enriched for extensin genes, which are associated with cell wall formation ($P = 1.41e–12$). The genes that have at least 10-fold higher expression in roots are significantly more methylated in shoots than roots (with 39% more sites with greater methylation in shoots than roots than the typical gene), and this effect is more pronounced at CG sites. As shown in **Figure 2B**, we see that the differentially methylated CG sites of these genes (normalized by total methylation levels) are strongly enriched around the transcription start site, again suggesting these regions have more variable methylation between these tissues. In contrast, we find that non-CG methylation appears to be relatively unchanged within the genes with higher expression in roots than shoots. These results indicate that methylation changes between roots and shoots are significantly enriched over genes with differential expression, suggesting that these sites may have a direct or indirect role in the regulation of these transcriptional changes.

Nucleosomes in shoots and roots

We identified nucleosome positions across our 2 samples by digesting chromatin with micrococcal nuclease and sequencing the mononucleosome fractions. These were mapped to the genome using the method described by Bowtie.[12] Each read in this library represents a putative nucleosome start site.

We first measured nucleosome density across the genome by computing the number of reads in 1000 kilobase windows. We find that nucleosome densities are enriched in the pericentromeric region compared with the arms (**Fig. 3A**). We also find that these nucleosome distributions differ slightly between root
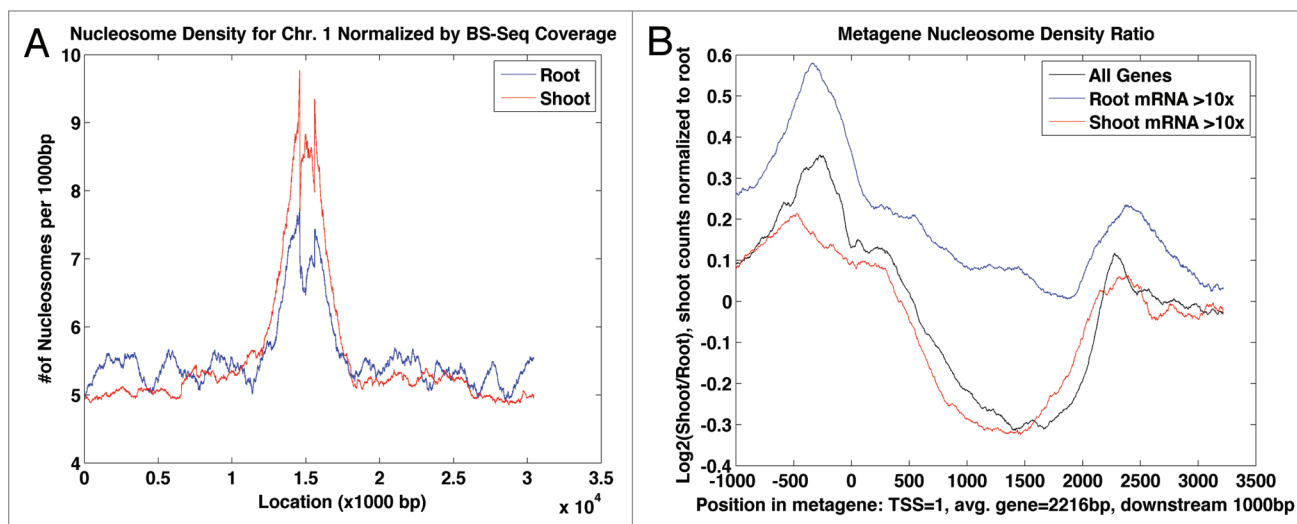
Figure 3. Nucleosome Density and Gene Expression. (**A**) Nucleosome density of roots (blue) and shoots (red). X, location on chromosome 1 (1000 bp increments); Y, nucleosome density: nucleosomes per 1000 bp with total number of nucleosomes in root and shoot normalized by BS-Seq coverage. (**B**) Average nucleosome density ratio for all genes (black), genes with 10× greater expression in shoots (red) and genes with 10× greater expression in roots (blue). X, location with respect to transcription start site, gene length normalized to 2216 bp; Y, log ratio (base 2) of shoot nucleosome density divided by root nucleosome density, shoot data normalized to root coverage level.

and shoot (**Fig. 3A**), with shoots having higher nucleosome densities in pericentromeric regions than roots.

We also measured the average nucleosome occupancy across genes (**Fig. 3B**). When we compared the nucleosome occupancy between shoots and roots we found that shoots had a higher nucleosome density before transcription start sites and had a second peak at the end of the gene, but were generally less nucleated over the gene body. This is a surprising result suggesting that there are global changes in genic chromatin structure between the 2 tissues with roots showing higher nucleosome densities over genes. Strikingly, genes that have at least 10-fold higher expression in root show considerably more nucleosomes around the transcriptional start and end sites in the shoot tissue samples, suggesting that increasing transcriptional rates decreases nucleosome densities over these regulatory regions. However, this effect was not seen for the shoot overexpressed genes, which have a nucleosome ratio profile that is very similar to that for all genes.

Nucleosome positioning and methylation

We measured methylation levels with respect to the start of nucleosomes and found, consistent with the results of our previous study,[8] that DNA methylation is enriched over the first nucleosome and that nucleosomes show a strong periodicity in methylation patterns that persist over several nucleosomes (**Fig. 4A**). This periodicity is approximately 180 base pairs long and is most pronounced in root tissue methylation. We performed a similar analysis of nucleosome periodicity by looking at the average density of nucleosome start sites centered on the start of each MNase read (**Fig. 4A**). We see a very similar periodicity of about 180 bases, as seen in the methylation. Finally, we asked if the 2 periodic patterns can be superimposed by shifting one with respect to the other. This allows us to determine where the peak of methylation occurs with respect to the peak of nucleosome positioning. To accomplish this we slid a 180 base window

of the methylation data across the nucleosome density data from 500 bases upstream to 500 bases downstream of a nucleosome start site. We find that the 2 patterns are maximally correlated when they are shifted by 25 bases, indicating that the peak of methylation occurs 25 bases before nucleosome start sites.

Extensin genes and transposons

We performed an overlap analysis on all differentially expressed, methylated, and nucleated genes and found the overlap to be statistically significant especially for genes that have higher expression in root as well as between nucleosome density and methylation level in both shoots and roots (**Fig. S1**).[13] Among the root overexpressed genes, we identified a small group that have both methylation, nucleosome density and gene expression changes between roots and shoots. These genes are significantly less methylated in roots compared with shoots, have at least a 10-fold expression preference in roots over shoots and have roughly one half the nucleosome density in roots than shoots (**Fig. 5**). Based on Gene Ontology terms, these genes were found to be in the extensin family of genes and are involved in cell wall formation.[14,15] Nine of the 10 extensin genes are found among the 11 genes that overlap all 3 differential sets of genes (**Table S2**). This suggests that this group of genes might play a previously unreported role in differentiating root cells from other cell types in the plant. The differences in methylation and nucleosome densities for two of these genes are shown in **Figure 5A and B**.

The methylation levels of several extensin genes were validated using traditional bisulfite sequencing and the results are reported in the supplementary material. With respect to CG methylation, the validation shows the strongest support for 4 of the 9 genes the validation was performed on. If we only consider only those cases for which both the BS-Seq data and the validation show a decrease or increase from shoot to root, 7 of these
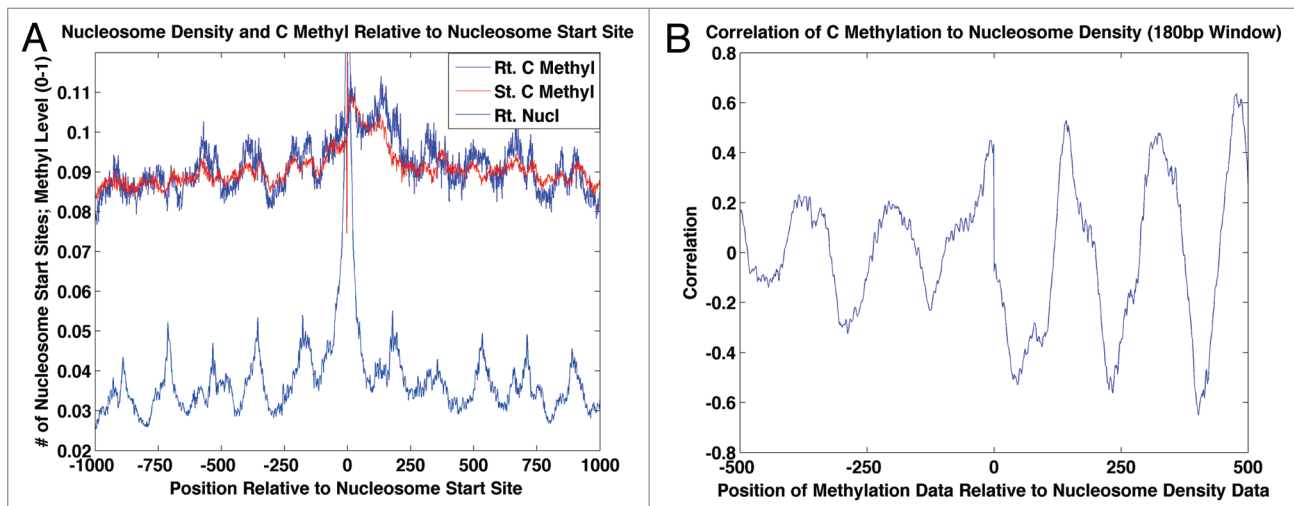
Figure 4. Methylation levels relative to nucleosome start site. (**A**) Nucleosome distribution relative to nucleosome start site compared with cytosine methylation. X, 1 kb upstream to 1 kb downstream; Y, # of nucleosomes, methylation level. (**B**) Correlation of methylation level to nucleosome density using a 180 base pair window. X, distance from start of methylation data window to start of nucleosome data window; −, upstream; +, downstream; Y, correlation.

genes have equivalent trends for CG methylation and 5 for either CHG or CHH methylation.

Transposons also show epigenetic differences between roots and shoots. Differential nucleosome density is found in numerous transposons with 11.6% of transposons showing at least a 2-fold difference and 37.4% showing at least a 50% difference. There are a total of 293 transposons with at least a 10% difference in CG methylation level with 82 having more methylation in shoots and 211 having more methylation in roots. Performing an overlap analysis on transposons found statistically significant results between differential methylation levels and nucleosome density in both shoots and roots but did not find any significance between expression and methylation or nucleosome density (**Fig. S1**). We also identified several transposons that showed significantly different epigenetic and transcriptional patterns between the 2 tissues. Some transposons that show differences in nucleosome density between shoot and root also have differences in expression levels. An example of a transposon with significant changes in methylation, nucleosome density and expression is shown in **Figure 5C**.

Discussion

We have shown that the root and shoot tissues from Arabidopsis accumulate significant epigenetic changes. We identify the largest changes in CG methylation in heterochromatic regions. Within genes, the transcriptional start (TSS) and end sites show the greatest variability in methylation when normalized by the average methylation levels. That is, although TSSs are generally hypomethylated, the levels of this hypomethylation are quite variable across these two tissues, indicating that these regions may be undergoing more chromatin changes that the bodies of genes. In contrast, within genes, our method does not have the resolution to identify significant changes in methylation at non-CG sites that are sparsely methylated. We have also identified clusters of significantly varying sites, and found that

these are enriched in extensin genes, that may represent a novel class of genes involved in root development, although no additional functional characterization is available for these genes.

We also noted a large number of loci with significantly altered chromatin structure, as measured by nucleosome positioning. Overall, the density of root nucleosomes in the centromeric regions of the chromosomes appears to be lower in roots than shoots, suggesting possible changes in the density of heterochromatin between these 2 tissues. We also found that genes that were root specific tended to be denucleated around the transcriptional start sites compared with shoots. Finally, in support of previous findings, we identify that both the positioning of nucleosomes and DNA methylation show a strong periodic behavior. We conclude that nucleosomes are preferentially methylated, and that there is a peak of DNA methylation about 25 bases before the nucleosome cut site, suggesting that the DNA that is adjacent to the linker DNA is most accessible to DNA methyltransferases.

Finally we show dramatic changes in DNA, gene expression and nucleosome structure around certain transposons, suggesting that these are often reactivated in a tissue specific manner. Future studies on specific cell types will undoubtedly find much more significant variation in gene expression, methylation, and chromatin structure than we could observe in our 2 coarse sectionings of plants into roots and shoots. Nonetheless even these results suggest that significant epigenetic changes accompany the transcriptional reprogramming that results from the development of roots and shoots.

Materials and Methods

All material was collected from the Columbia strain of *Arabidopsis thaliana*. We used the genome annotation from TAIR7.
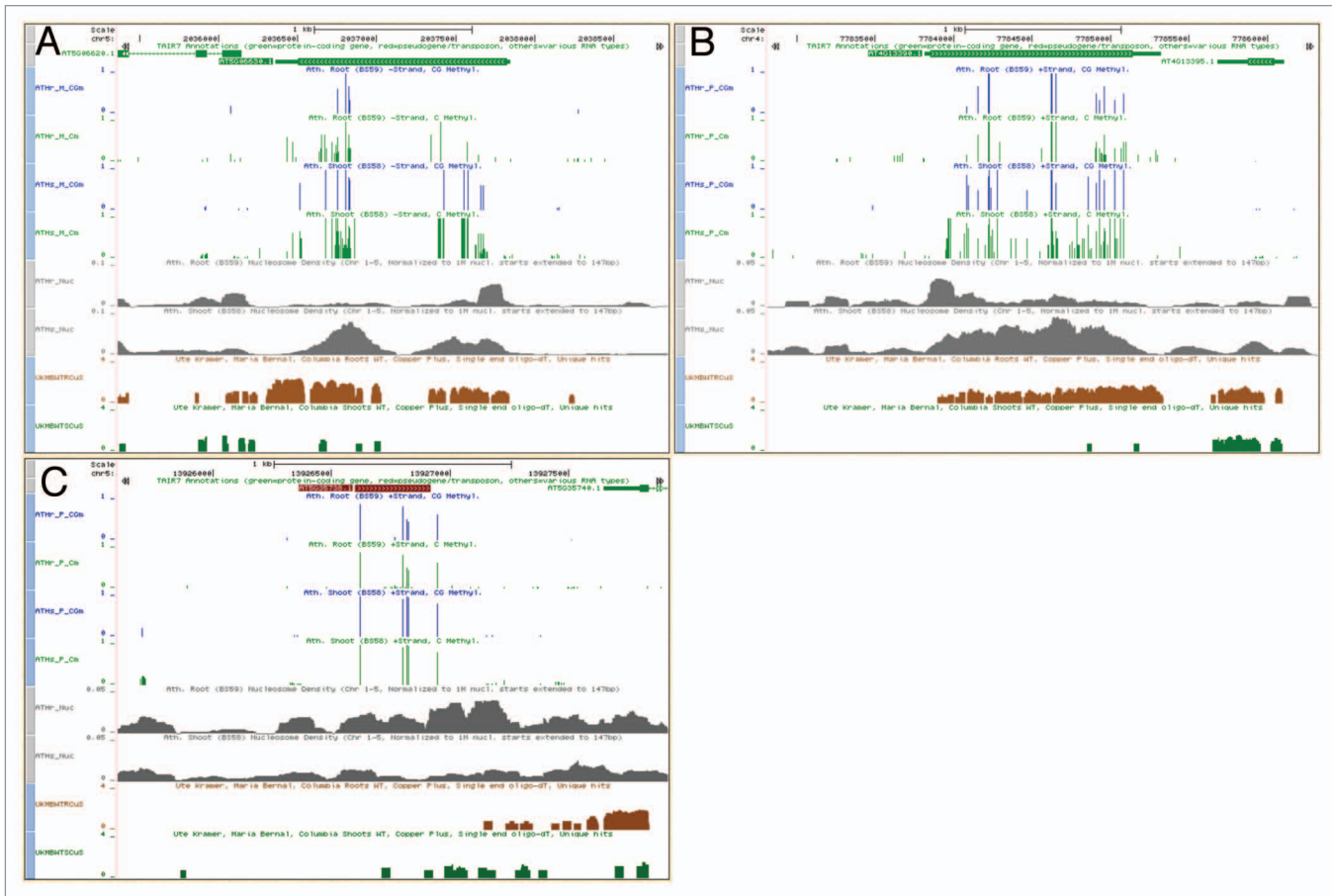
Figure 5. Comparison Between Root and Shoot in UCSC Genome Browser. (**A and B**) Select Extensin Genes: gene on -strand on left (**A**), gene on +strand on right (**B**). (**C**) Example of a transposon/pseudogene. Transposons that had a significant difference in nucleosome density between root and shoot showed the most difference in expression level. Blue histograms: CG methylation on same strand as genes in roots(top)/shoots(bottom). Green histograms: C methylation on same strand as genes in roots(top)/shoots(bottom). Gray histograms: Nucleosome density in roots(top)/shoots(bottom). Brown and dark green histograms: Density of mapped mRNA in roots(brown, top)/shoots(dark green, bottom).

Bisulfite libraries and analysis

For the methylation analysis we used bisulfite sequencing of whole genome libraries following previously published protocols utilizing pre-methylated adapters.[9] The 2 libraries, root and shoot, were prepared from roots and all tissues of the plant growing above ground, respectively. Samples for the root library were from a root culture and are not from the same plants that were used for the shoot library. The roots were obtained from one-month-old Arabidopsis grown in Gamborg's B5 liquid medium under continuous white light. After obtaining the raw BS-seq reads, the reads were mapped using the method described by Seeker. Each base that is mapped to a cytosine can either be a cytosine or converted to a thymine by the bisulfite reagents. Methylated cytosines remain as cytosines after bisulfite treatment but unmethylated cytosines are deaminated into uracil, which is reverse transcribed into thymine in our reads. The methylation percentage at each site is determined by measuring the ratio of the cytosines mapped to a particular location over the total number of reads. We measured single-site differential methylation by comparing sites with an absolute difference of 30% methylation (e.g., 20% vs 50%, not relative difference such

as 20% vs 26%). In order make sure that these differential sites were statistically significant only sites that had a binomial distribution confidence interval of 5% were included. The confidence interval was calculated by computing the binomial cumulative distribution function of the BS-seq data at each site that had an absolute methylation difference above the threshold of 30% and only incorporated sites into the single site differential data where the BS-seq data scored less than 0.025 or greater than 0.975 in the CDF.

RNA-seq

We sequenced and mapped mRNA to determine gene expression levels. The expression level was measured by normalizing the number of times mRNA mapped to a particular gene in the reference genome to a million mRNA reads. This reports the data in terms of reference sequence hits per million reads. As these libraries were oligo DT primed, they do not cover the entire transcript and we therefore did not normalize by transcript length.

Nucleosomes

Nucleosome data was obtained by sequencing MNase digested DNA. MNase cleaves DNA at both ends of nucleosomes leaving any DNA bound to the histone core of the nucleosome

unaffected. This nucleosome-bound DNA was then separated from the histones and sequenced.[8] This sequenced DNA was mapped to the reference genome using the method described by Bowtie,[12] allowing for 2 mismatches and the nucleosome start sites were determined by looking at the first position of each mapped read.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Supplemental Materials

Supplemental materials may be found here:
www.landesbioscience.com/journals/epigenetics/article/26869

References

1. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 2003; 33(Suppl):245-54; PMID:12610534; http://dx.doi.org/10.1038/ng1089

2. Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 2002; 297:1833-7; PMID:12193640; http://dx.doi.org/10.1126/science.1074973

3. Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, et al. Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci U S A 2010; 107:8689-94; PMID:20395551; http://dx.doi.org/10.1073/pnas.1002720107

4. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 2008; 452:215-9; PMID:18278030; http://dx.doi.org/10.1038/nature06745

5. Hsieh TF, Ibarra CA, Silva P, Zemach A, Eshed-Williams L, Fischer RL, Zilberman D. Genome-wide demethylation of Arabidopsis endosperm. Science 2009; 324:1451-4; PMID:19520962; http://dx.doi.org/10.1126/science.1172417

6. Song F, Mahmood S, Ghosh S, Liang P, Smiraglia DJ, Nagase H, Held WA. Tissue specific differentially methylated regions (TDMR): Changes in DNA methylation during development. Genomics 2009; 93:130-9; PMID:18952162; http://dx.doi.org/10.1016/j.ygeno.2008.09.003

7. Zemach A, Kim MY, Silva P, Rodrigues JA, Dotson B, Brooks MD, Zilberman D. Local DNA hypomethylation activates genes in rice endosperm. Proc Natl Acad Sci U S A 2010; 107:18729-34; PMID:20937895; http://dx.doi.org/10.1073/pnas.1009695107

8. Chodavarapu RK, Feng S, Bernatavichute YV, Chen PY, Stroud H, Yu Y, Hetzel JA, Kuo F, Kim J, Cokus SJ, et al. Relationship between nucleosome positioning and DNA methylation. Nature 2010; 466:388-92; PMID:20512117; http://dx.doi.org/10.1038/nature09147

9. Feng S, Rubbi L, Jacobsen SE, Pellegrini M. Determining DNA methylation profiles using sequencing. Methods Mol Biol 2011; 733:223-38; PMID:21431774; http://dx.doi.org/10.1007/978-1-61779-089-8_16

10. Chen PY, Cokus SJ, Pellegrini MBS. BS Seeker: precise mapping for bisulfite sequencing. BMC Bioinformatics 2010; 11:203; PMID:20416082; http://dx.doi.org/10.1186/1471-2105-11-203

11. Bernal M, Casero D, Singh V, Wilson GT, Grande A, Yang H, Dodani SC, Pellegrini M, Huijser P, Connolly EL, et al. Transcriptome sequencing identifies SPL7-regulated copper acquisition genes FRO4/FRO5 and the copper dependence of iron homeostasis in Arabidopsis. Plant Cell 2012; 24:738-61; PMID:22374396; http://dx.doi.org/10.1105/tpc.111.090431

12. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009; 10:R25; PMID:19261174; http://dx.doi.org/10.1186/gb-2009-10-3-r25

13. Oliveros JC. An interactive tool for comparing lists with Venn Diagrams. 2007

14. Neubauer JD, Lulai EC, Thompson AL, Suttle JC, Bolton MD. Wounding coordinately induces cell wall protein, cell cycle and pectin methyl esterase genes involved in tuber closing layer and wound periderm development. J Plant Physiol 2012; 169:586-95; PMID:22251796; http://dx.doi.org/10.1016/j.jplph.2011.12.010

15. Velasquez SM, Ricardi MM, Dorosz JG, Fernandez PV, Nadra AD, Pol-Fachin L, Egelund J, Gille S, Harholt J, Ciancia M, et al. O-glycosylated cell wall proteins are essential in root hair growth. Science 2011; 332:1401-3; PMID:21680836; http://dx.doi.org/10.1126/science.1206657