Dynamics of the epigenetic landscape during development and in response to drought stress in sorghum

Yongfeng Hu[1, †, *], Chao He[2, †], Yuning Shen[2], Gongjian Zeng[1], Siteng Bi[2], Quanjun Huang[1], Xiner Qin[1], Zhuying Deng[1], Zhengquan He[1], Xiangling Shen[1,*]

[1]Key Laboratory of Three Gorges Regional Plant Genetics & Germplasm Enhancement (CTGU)/ Biotechnology Research Center, College of Biological and Pharmaceutical Sciences, China Three Gorges University, Yichang, Hubei 443002, China.

[2]National Key Laboratory of Crop Genetic Improvement, Hubei Hongshan Laboratory, Huazhong Agricultural University, Wuhan, China.

[†]These authors contributed equally to this work.

[*]Corresponding author: huyongfeng@ctgu.edu.cn; shenxl1982@hotmail.com

Running title: Sorghum epigenome.

Abstract

Background: *Sorghum bicolor* is a C4 plant with the characteristics of high stress tolerance, which may be conferred partly by the underlying epigenetic mechanism unique to sorghum.

Results: In this study, we investigated the differential enrichment of 7 histone marks in leaf and root as well as in response to PEG treatment in the two organs, and revealed some epigenomic features that have never been reported before. For example, the long H3K27me3 regions clustered in four areas, which we defined as H3K27me3 islands, were identified in sorghum. The increase of some H3K27me3 enrichment was associated with gene up-regulation in leaf but not in root. H3K36me3 plays some role in inhibiting the deposition of both H3K27me3 and H2A.Z, which may serve as partial motivation for the removal of H3K27me3 and H2A.Z in leaf and root, but was not involved in the prevention of H3K27me3 spreading into a long region. Finally, we analyzed histone mark variation in the genes that were differentially expressed both between leaf and root and in response to PEG treatment (defined as common DEGs), and found that most of the changes of histone marks occurred in some common DEGs in leaf and root could not be observed in the same genes in response to PEG treatment.

Conclusions: These findings provide useful information for further studying the role of epigenetic regulation in sorghum development and stress response as well as new clues for the analysis of epigenetic mechanism in plants.

Key words: sorghum, histone modification, H2A.Z, root and leaf, drought stress

Background

Histone marks play important roles in gene regulation, and they exhibit different distribution patterns in intragenic regions. Histone acetylation and H3K4me3 tend to mark transcription start site (TSS) regions (a short region immediately downstream from TSS), while H3K36me3 are more likely to be enriched in gene body regions (a longer region from TSS to transcription end site (TES)) [1-3], suggesting that they are involved in governing different processes of gene transcription. Histone acetylation and H3K4me3 have been reported to be involved in transcription initiation by affecting the assembly of pre-initiation complex (PIC), or stimulating the transition from initiation to elongation [4-7]. H3K36me3, generally considered as a hallmark of transcription elongation, is required for co-transcriptional splicing [2, 8, 9]. The distribution patterns of some histone marks are similar in different species while the others are not. For example, in humans and yeast, H3K36me3 is excluded from TSS of genes while there is a clear enrichment peak of the mark around TSS in plants [1, 10-12]. In animals, H3K27me3 can spread to hundreds of kilobases and cover several genes per region [13]. However, most of H3K27me3 regions in *Arabidopsis* are limited to single genes [14]. The different distribution patterns of these histone marks may reflect diverse deposition mechanism of them in different species.

The complex crosstalk between different histone modifications is established, since the activity of histone modifiers depositing one histone mark could be either promoted or inhibited by other histone marks. For instance, H2B ubiquitylation is required for simulating the activity of the MLL complex to catalyze H3K4me3 [15]. H3K4me3 provides the binding platform for the HAT complexes, facilitating histone acetylations [16, 17]. The interplay between repressive marks such as H3K27me3 and H2AK119ub1, which are deposited by Polycomb repressive complex 2 (PRC2) and Polycomb repressive complex 1 (PRC1) respectively, have also been unraveled [18]. Interestingly, both PRC2 and PRC1 can read the two marks to establish or reinforce these modifications [18]. By contrast, H3K4me3 and H3K36me3 inhibit the activity of PRC2 [19, 20]. Furthermore, it has been demonstrated that H3K36me3 and H3K27me3 are rarely present on the same histone tail in various species [21, 22], suggesting that the inhibitory effect of H3K36me3 on deposition of H3K27me3 is conserved. Indeed, genome-wide analysis of chromatin states indicates that the distributed regions of some histone marks highly co-exist while the others are seldom overlapped, which is well consistent with the crosstalk between them [23, 24].

Sorghum is an ideal model species for analyzing genetic features of C4 plants due to its small genome and natural diploid. Besides, it also exhibits high tolerance to various abiotic stresses, making it a good candidate for studying mechanism of plant stress tolerance. In this paper, genome-wide analysis of various histone modification enrichment was conducted in sorghum root and leave before and after PEG treatment to explore the role of epigenetic marks in the regulation of sorghum development and stress response. We found that distribution pattern of these marks in sorghum was similar to that reported in other plant species. However, sorghum-specific epigenomic features were also revealed. For example, four H3K27me3 islands were identified where long H3K27me3 regions clustered, which may be involved in shaping specific chromatin architecture in sorghum. In leaf, specific intragenic H3K27me3 enrichment was positively associated with gene expression, which was not observed in root. H3K36me3 enrichment inhibited deposition of both H3K27me3 and H2A.Z,

but not played a role in preventing H3K27me3 spreading. In addition, for many genes that were both differentially expressed between leaf and root and activated or repressed by PEG treatment, variation in histone mark enrichment occurred only between the two organs but not in response to stress.

Results

Obtainment of transcriptomic and epigenomic data from sorghum root and leaf

To obtain the transcriptomic and epigenomic data of sorghum, two-week-old seedlings of BTx623 variety were treated with 20% PEG for 6 hours before roots and leaves were harvested for RNA-seq as well as ChIP-seq with antibodies against various histone modifications (H3K9ac, H3K27ac, H3K4me3, H3K36me3, H3K4me2 and H3K27me3) and H2A.Z. Three replicates for RNA-seq and two biological replicates for ChIP-seq were performed. The three replicates of RNA-seq analysis were highly correlated (Additional file 1: Fig. S1A). For most of the ChIP-seq samples, more than 20 million reads were obtained, the overall mapping rate was more than 85% (Additional file 1: Table S1), the FRiP value and the SPOT value were greater than 0.5 (Additional file 1: Fig. S1B). The two biological replicates of ChIP-seq were also highly correlated (Additional file 1: Fig. S1C). Genome-wide view of histone mark enrichment demonstrate that the distribution of all the 7 histone marks was highly correlated with gene density, which is relatively higher at both ends of each chromosome and lower in the middle of the chromosome (Additional file 1: Fig. S2). These results suggest that all the data are reliable for subsequent analysis.

Distribution pattern of histone marks in sorghum

To attempt to investigate the distribution pattern of histone marks, intragenic peaks (located from 1.5 kb upstream from gene TSS to 1.5 kb downstream from gene TES) were called for creating metaplots and heat maps (Fig. 1A and 1B). The results showed that H3K9ac, H3K27ac, H3K4me3 were predominantly enriched in TSS regions. H3K4me2 and H2A.Z were distributed either TSS regions or gene body regions. H3K36me3 enrichment extended from TSS in gene body but not to TES. The distribution patterns of these histone marks are similar to those in *Arabidopsis* and rice [3, 10, 25, 26]. We also found strong H3K27me3 enrichment in gene body in both root and leaf. Interestingly, low H3K27me3 enrichment in TSS regions was observed in leaf but not in root (Fig. 1B), which has never been reported in plants before. Besides, genome-wide enrichment of H3K4me2, H3K36me3, H3K27me3 and H2A.Z was decreased in root, suggesting their possible regulatory roles in the fine-tuning of developmental genes (Fig. 1A). To explore the genome-wide nucleosome profile in sorghum, MNase-seq was performed using the same materials as ChIP-seq. We found that different from five or more phased nucleosomes in rice and *Arabidopsis* [27, 28], there were only three peaks representing clearly phased nucleosomes (numbered as +1, +2 and +3) downstream from the TSS in both leaves and roots of sorghum (Fig. 1C and 1D). However, the peak indicating -1 nucleosome was not clear upstream from the TSS, although nucleosome free region (NFR) did apparently exist (Fig. 1C and 1D). In other words, the length of NFR is highly variable among different genes in sorghum. These results demonstrate that sorghum epigenome exhibits both conserved and unique features in comparison to other plant species.

In animals, H3K27me3 enrichment can spread to more than one hundred kilobases,

which usually cover several genes [13]. It seems that in *Arabidopsis* H3K27me3 coverage is confined to single gene [14]. To determine whether H3K27me3 in sorghum is distributed similarly, we analyzed length of H3K27me3 regions and number of genes per region. The results showed that most of H3K27me3 regions were short and covered single genes in sorghum (Fig. 2A and 2B), which is similar as in *Arabidopsis*. However, we found considerable long H3K27me3 regions in sorghum (Fig. 2A and 2B). The maximum length can reach more than two hundred kilobases. Three H3K27me3 regions covered more than 20 genes (Fig. 2B). In addition, there were more short regions and less long regions in leaf than those in root, suggesting that H3K27me3 enrichment in the two organs are reprogrammed. When we visualized genome-wide H3K27me3 regions, we found four areas on Chromosomes 2, 5 and 8 where long H3K27me3 regions clustered, which we termed the H3K27me3 islands (Fig. 2C). The other long H3K27me3 regions scattered in the other part of the genome. The similar analysis was performed in *Arabidopsis* and rice by using the recently deposited high quality data (see Material and methods). The long H3K27me3 regions were also identified in *Arabidopsis* and rice despite that the number was less than that in sorghum (Additional file 1: Fig. S3A), especially the regions longer than 50kb. In addition, the H3K27me3 island was not observed in *Arabidopsis* and rice (Additional file 1: Fig. S3C). However, the genes covered by long H3K27me3 regions were greater in *Arabidopsis* (Additional file 1: Fig. S3B), which may be due to the higher gene density. This indicates that plant epigenomes contain both gene-specific and long region of H3K27me3 enrichment.

Crosstalk between different histone marks

The deposition of different histone marks can be either promoted or inhibited by each other, which is reflected by their overlapping or exclusive distributions. To analyze the relationship between different histone marks in sorghum, we clustered distribution patterns of each histone mark in intragenic region (from 0.5kb upstream TSS to 3kb downstream TSS) into three categories (C1, C2 and C3), and observe how the other histone marks are distributed in the same category. We found that H3K9ac, H3K27ac, H3K4me3 enriched in the TSS region and H3K36me3 enriched in the gene body region are highly overlapped and associated with actively transcribed genes (Fig. 3A-3D, S4A-S4D). The gene-body- enriched H3K27me3 is exclusive with the above histone modifications especially H3K36me3, and is associated with silent genes (Fig. 3G, Additional file 1: Fig. S4G). The partial overlapping of H3K4me3 and H3K27me3 suggests the existence of the bivalent mark in sorghum. The distributions of H3K4me2 and H2A.Z in different regions displayed distinct relationships with the other marks (Fig. 3E-3F, Additional file 1: Fig. S4E-S4F). The strong enrichment of H2A.Z in the gene body region (C1) overlapped with the repressive mark (H3K27me3), whereas the weak enrichment of H2A.Z in the TSS region (C3) overlapped with the active marks (H3K9ac, H3K27ac, H3K36me3) (Fig. 3E, Additional file 1: Fig. S4E). By contrast, H3K4me2 deposited in the TSS region (C1) co-existed with the repressive mark, whereas that deposited in the gene body region (C2) co-existed with the active marks (Fig. 3F, Additional file 1: Fig. S4F). Furthermore, we found that H3K4me2 and H3K4me3 co-existed in many genes although distribution patterns of the two marks in these genes were slightly different. This suggests that the nucleosome simultaneously carrying H3K4me2 and H3K4me3 is common in sorghum, but the deposition of the two marks may be independent.

To further analyzing the crosstalk among different histone marks, we identified 8

chromatin states (CS1-CS8) in sorghum genome by combining different histone marks (Additional file 1: Fig. S5A). The expression level of genes in different chromatin states was examined (Additional file 1: Fig. S5C). The genes in CS2, CS4 and CS5 were expressed in relatively lower levels than those in the other chromatin states, indicating that they represent the repressive chromatin states (Additional file 1: Fig. S5C). CS5 was associated with only H3K27me3 while in CS4 both H2A.Z and H3K27me3 were enriched (Additional file 1: Fig. S5A). H3K4me2 and H2A.Z were abundant in CS2 that was also associated with less H3K4me3 and H3K27me3 (Additional file 1: Fig. S5A). The genes with relatively higher expression levels resided in CS1 marked by H3K36me3 and H3K4me2, CS7 marked by H3K9ac, H3K27ac and H2A.Z and CS8 marked by all the 7 histone marks (Additional file 1: Fig. S5A, S5C). Surprisingly, CS6 deficient in any of the 7 histone marks occupied most of the genome, and covered also a great number of actively transcribed genes comparable to expression level of genes in CS3 where only H2A.Z was prevalent (Additional file 1: Fig. S5A-S5C). In addition, analysis of genomic annotation enrichment indicated that the proportion of genes in CS1, CS2, CS3, CS4 and CS8 was larger than that in the other chromatin states (Additional file 1: Fig. S5B).

Considering the strictly antagonistic distribution of H3K27me3 and H3K36me3, we wonder how H3K36me3 is distributed in H3K27me3 islands. We selected one area in each H3K27me3 island to visualize the distribution of H3K27me3 and H3K36me3. The diagram showed that H3K27me3 and H3K36me3 seldom co-existed at the same gene (Fig. 3H, Additional file 1: Fig. S4H), suggesting that they may exclude each other. Indeed, the long H3K27me3 regions are disrupted due to the deposition of H3K36me3. Thus, we speculate that there should be much more and longer H3K27me3 regions than expected in sorghum. As H2A.Z deposition can also extend to generate long H2A.Z regions (data not shown) and H2A.Z is absent in the gene body regions marked by H3K36me3 (Fig. 3E), we randomly selected three long H2A.Z regions and we found that enrichment of H3K36me3 excluded H2A.Z in these regions (Fig. 3I, Additional file 1: Fig. S4I). However, the antagonistic relationship between H3K36me3 and H2A.Z is not as rigorous as H3K36me3 and H3K27me3, as at some loci H3K36me3 and H2A.Z co-existed. Taken together, these data suggest that H3K36me3 inhibits the deposition of both H3K27me3 and H2A.Z. To confirm whether the scenario is conserved in plants, we performed the similar analysis in *Arabidopsis* and rice. The results also revealed the antagonistic distributions between H3K27me3 and H3K36me3 as well as H2A.Z and H3K36me3 (Additional file 1: Fig. S5A-D).

Dynamic change of histone marks during development in sorghum

In order to identify which histone marks are more likely to be related with specific gene expression during developmental processes, correlation analysis of differential enrichment of histone marks and gene expression in leaf and root was performed. The metaplots indicated that the overall levels of H3K9ac, H3K27ac, H3K4me3 and H3K36me3 were increased both at root-up-DEGs (Gene expression is up-regulated in roots compared to leaves) in root and leaf-up-DEGs (Gene expression is up-regulated in leaves compared to roots) in leaf (Fig. 4A). The number of genes whose expression was positively correlated with changes in these marks was much greater than the number of genes whose expression was negatively correlated with them (Fig. 4B). These supports the positive role of these marks in the regulation of gene expression. In contrast, H2A.Z, H3K4me2 and H3K37me3 levels were significantly decreased at

root-up-DEGs in root but slightly decreased at leaf-up-DEGs in leaf (Fig. 4A), and genes negatively regulated by these marks were more than those positively regulated (Fig. 4B), suggesting negative correlation of these marks with gene expression.

To investigate how different histone marks cooperate to regulate root-leaf-DEGs (differentially expressed genes between root and leaf), we calculated the proportions of genes that were co-regulated by two histone marks. Considering the opposite roles of different histone marks in gene regulation as described above, the number of genes with increased H3K9ac, H3K27ac, H3K4me3, H3K36me3 or decreased H2A.Z, H3K4me2, H3K27me3 between leaf and root were used for the calculation. The higher proportions of overlapping (more than 30%) were observed between the increase of H3K9ac and H3K27ac as well as the increase of H3K4me3 and H3K36me3 in both leaf (vs root) and root (vs leaf) (Fig. 4C), suggesting that these two pairs of histone marks tend to activate gene expression collaboratively. The decrease of H3K4me2 and H3K27me3 also showed a high percentage (31.13%) of overlapping only in leaf (vs root) (Fig. 4C), which demonstrates that they cooperate to repress more gene expression in root. We found that the proportions of overlapping between the increase of H3K36me3 and the decrease of H3K27me3 as well as the decrease of H2A.Z deposition were not as high as expected (8.55% and 15.65% in leaf (vs root), 8.11% and 9.80% root (vs leaf)) (Fig. 4C). This indicates that H3K36me3 enrichment contribute to small part of decline of H3K27me3 enrichment or H2A.Z deposition. In addition, the heat map showed that nearly half of genes with increased H3K36me3 were not marked with H3K27me3 (Fig. 4D). However, most of strong H3K27me3 enrichment were inhibited by the increase of H3K36me3. Similarly, the relatively higher levels of H2A.Z deposition were decreased to moderate levels after H3K36me3 enrichment was increased (Fig. 4E). These results further support the notion that H3K36me3 play part of the roles in the inhibition of H3K27me3 enrichment or H2A.Z deposition.

There were 7006 root-up-DEGs and 5100 leaf-up-DEGs which comprises about one third of total sorghum coding genes (n=34,118). We would like to learn how many root-leaf-DEGs involve the dynamic change of histone marks. The percentages of genes with differential enrichment of 0 to 7 histone marks were calculated for root-up-DEGs and leaf-up-DEGs respectively. The results indicated that 44.39% of root-up-DEGs and 29.92% of leaf-up-DEGs were not related to changes of any histone marks (Additional file 1: Fig. S7A), and 95% and 85% of these genes carried 1 to 7 histone marks (Additional file 1: Fig. S7B). In other words, it is not possible that 7 histone marks contribute to differential expression of these genes between root and leaf. Furthermore, by analyzing enriched levels of each histone marks in these genes, we observed that only small proportions of these genes were marked with H3K27me3, while majority of them were marked with the other marks and in moderate levels mostly (Additional file 1: Fig. S7C).

The H3K27me3 enrichment peak in TSS region of leaf-up-DEGs in leaf is much sharper than that in root (Fig. 4A). The heat maps in Fig. 1B have shown that many genes with TSS-enriched H3K27me3 was specific in leaf (Fig. 1B). There were considerable amounts of leaf-up-DEGs with increased H3K27me3 level in leaf (Fig. 4B). These results prompt us to suspect that TSS-enriched H3K27me3 in leaf may be positively associated with gene expression. H3K27me3 levels in the 456 genes with increased gene expression and H3K27me3 enrichment in leaf (vs root) and in the rest of leaf-up-DEGs other than these

genes were investigated. The metaplots showed that a clear sharp peak of H3K27me3 enrichment in TSS region of the 456 genes in leaf but not in root, and the increase of H3K27me3 in leaf confined to the gene body region (Additional file 1: Fig. S8A). The peak of H3K27me3 enrichment in TSS region of the rest of leaf-up-DEGs in leaf (Additional file 1: Fig. S8B) was less sharp than that of all the leaf-up-DEGs (Fig. 4A), and H3K27me3 level in the rest of leaf-up-DEGs in leaf was significantly decreased than that in root (Additional file 1: Fig. S8B). These demonstrate that the sharp peak of H3K27me3 enrichment in TSS region of leaf-up-DEGs in leaf is partly due to the increase of H3K27me3 level in the 456 genes. By contrast, the gene-body-specific increase of H3K27me3 level in the 241 genes with increased gene expression and H3K27me3 enrichment in root (vs leaf) was not observed (Additional file 1: Fig. S8C). Instead, H3K27me3 enrichment upstream from TSS and downstream from TES was also increased (Additional file 1: Fig. S8C), suggesting that the increase of H3K27me3 enrichment might be due to the variations in long H3K27me3 regions covering multiple genes. Besides, the difference of H3K27me3 levels in the rest of root-up-DEGs other than the 241 genes was similar to that in all the root-up-DEGs (Additional file 1: Fig. S8D). However, the other histone marks are also possibly responsible for the regulation of these genes, as we found that expression of nearly half of the 456 genes and quarter of the 241 genes were positively associated with H3K9ac, H3K27ac, H3K4me3, or H3K36me3 (Additional file 1: Fig. S8E and S8F).

As shown in Fig. 4A, the differential enrichment of H3K27me3 in root-up-DEGs and leaf-up-DEGs in root and leaf was also obvious outside the gene body region (upstream of TSS and downstream of TES), suggesting that the dynamic change of H3K27me3 could spread outside the intragenic region. Alternatively, H3K27me3 deposition changes in units of long segments instead of single genes. To confirm the suspicion, we examined the number of H3K27m3 peaks with different length differentially enriched in root and leaf and found that although most of differentially enriched H3K27me3 peaks were short (less than 10 kb), more than 300 long H3K27me3 regions (more than 10 kb), consisting of 278 (10-20 kb), 87 (20-40 kb), 12 (40-60 kb), 5 (60-80 kb) and 4 (80-100 kb), were also differentially deposited (Additional file 1: Fig. S9A). Genome-wide screenshots indicated that the changes of long H3K27me3 regions scattered in 10 chromosomes and negatively co-related with gene expression in general (Additional file 1: Fig. S9B). But close observation on individual regions showed that a number of genes covered by long H3K27me3 regions were not activated despite removal of H3K27me3 (Additional file 1: Fig. S9B). For example, two neighboring long H3K27me3 regions decreased in root and localized in chr5:70,880 kb-71,000 kb cover 11 genes, among which only 3 genes were up-regulated (Additional file 1: Fig. S9B). This suggests that the alteration of H3K27me3 enrichment in long regions is not always associated with transcriptional activity.

Drought-responsive genome-wide variation of histone marks

To explore how histone marks engage in the induction of drought-responsive genes, we investigated the levels of histone marks in PEG-up-DEGs and PEG-down-DEGs in response to PEG treatment in root and leaf. The metaplots showed that PEG treatment only slightly affects the overall levels of some histone marks (Fig. 5A and 5B). For example, H3K9ac, H3K27ac, H3K4me3, H3K36me3 levels in PEG-up-DEGs in leaf and H3K36me3 level in PEG-up-DEGs in root showed a little increase. H3K9ac and H3K27ac levels in

PEG-down-DEGs in leaf and root was slightly decreased. H2A.Z and H3K4me3 levels in PEG-down-DEGs in leaf and root rose and H2A.Z level in PEG-up-DEGs in leaf and root declined. These results are consistent with the regulatory roles of these marks. Then we performed association analysis of changes of histone marks and gene expression. In both leaf and root, only a few genes had altered H3K36me3 and H3K27me3 responding to PEG treatment (Fig. 5C and 5D), which suggests that these two marks are probably not involved in the PEG-responsive gene expression regulation. H3K4me3 was more likely to be associated with gene activation in response to PEG treatment, as the number of genes with increased H3K4me3 levels and gene expression in leaf (n = 138) and root (n = 363) was far more than the number of genes with the other relationships between H3K4me3 and gene expression (Fig. 5C and 5D). H3K9ac and H3K27ac were associated with either gene activation or repression when the level of these marks was increased or decreased responding to PEG treatment in leaf and root (Fig. 5C and 5D). Surprisingly, for certain number of genes in root but few in leaf changes of H3K9ac and H3K27ac were negatively correlated with changes of gene expression, which seems to contradict to the positive role of these marks. Finally, H2A.Z and H3K4me2 were negatively associated with some gene activation or repression in leaf and root (Fig. 5C and 5D). These results demonstrate that only a small proportion of gene induction in response to PEG treatment require dynamic change of histone marks.

In order to find out whether histone marks affect the degree of gene induction by PEG treatment, we classified PEG-up-DEGs into three groups: genes without enrichment of histone marks, genes with unchanged histone marks after PEG treatment, genes with changed histone marks after PEG treatment, and investigated fold change of each group gene expression. As PEG-responding gene up-regulation are more associated with increase of H3K9ac, H3K27ac, H3K4me3 and decrease of H2A.Z, the effect of these marks on gene induction was analyzed. We found that the degree of induction of the genes with changed histone marks after PEG treatment was significantly higher than that of the other groups (Additional file 1: Fig. S10), suggesting that the dynamic change of histone marks can elevate gene induction to some extent. However, the genes with unchanged H3K9ac, H3K27ac and H3K4me3 were induced to a relatively lower level than the genes without histone marks (Additional file 1: Fig. S10). Conversely, the genes with unchanged H2A.Z enrichment were induced to a higher level (Additional file 1: Fig. S10). These results demonstrate that the basic level of histone marks is not related to the gene responsiveness.

Developmentally regulated change of histone marks rarely overlaps with that induced by PEG treatment

Our results have revealed many genes for which alteration of gene expression is constantly accompanied with change of histone marks between leaf and root or responding to PEG treatment. Therefore, we would like to learn that for the genes both differentially expressed between leaf and root and responsive to PEG treatment whether the histone marks also change in both conditions (developmentally regulated and PEG responsive). Firstly, we identified 780 and 2825 Root-up-DEGs that were also PEG-up-DEGs in leaf and PEG-down-DEGs in root respectively, 1359 and 950 Leaf-up-DEGs that were also PEG-down-DEGs in leaf and PEG-up-DEGs in root respectively (Fig. 6A), which were defined as four groups of common DEGs. Then we investigated the number of genes where histone marks changed uniquely in each condition or in both conditions among each of four

groups of common DEGs. For the up-regulated genes increase of H3K9ac, H3K27me3, H3K4me3, H3K36me3 and decrease of H2A.Z, H3K4me2, H3K27me3 were analyzed, and for the down-regulated genes decrease of H3K9ac, H3K27me3, H3K4me3, H3K36me3 and increase of H2A.Z, H3K4me2, H3K27me3 were analyzed. The results revealed that for all four groups of common DEGs, the number of genes with change of histone marks uniquely regulated developmentally was much greater than those uniquely responsive to PEG and those in both conditions (Fig. 6B). Moreover, we found some common DEGs where 7 histone marks changed developmentally but none of them changed in response to PEG (Additional file 1: Fig. S11). This indicates that the dynamic change of histone mark important for developmental regulation of gene expression is not necessary for the gene induction in response to environmental cues, and some specific development-specific regulated factors are possibly required for the alteration of histone marks.

Dynamics of nucleosome occupancy in response to PEG treatment in leaf and root

Analysis of genome-wide nucleosome profiles in response to PEG in leaf and root revealed that nucleosome occupancy from TSS to downstream 1kb in leaf was generally elevated compared to that in root, and PEG treatment lead to general decrease of nucleosome occupancy in leaf (Fig. 7A). The elevation of nucleosome occupancy was more prominent in Root-up-DEGs but not in Leaf-up-DEGs (Fig. 7B), suggesting that gene up-regulation in root or down-regulation in leaf involves nucleosome removal or assembly. The decrease of nucleosome occupancy in PEG-up-DEGs in leaf was significant, but no difference was found in PEG-down-DEGs in leaf as well as in PEG-up-DEGs and PEG-down-DEGs in root (Fig. 7C and 7D). This indicates that induction of genes in leaf also require nucleosome rearrangement. To ascertain whether dynamics of nucleosome occupancy in response to PEG is related to change of histone marks, we examined the nucleosome profiles in the genes with changes of H3K9ac, H3K27ac, H3K4me3 and H2A.Z in leaf and root. The phased nucleosomes could not be observed in all the genes with changed histone marks in leaf, but the nucleosome level was clearly decreased in the genes with increased H3K9ac and HK27ac (Fig. 7E). In contrast, the +1 nucleosome was prominent in the genes with changed histone marks in root, and the three phased nucleosomes were clear in the genes with increased H3K9ac, decreased H3K9ac, increased H3K27ac, increased H2A.Z (Fig. 7F). But the variation of nucleosome occupancy was not obvious except in the genes with altered H2A.Z (Fig. 7F). Interestingly, we found that in the genes with increased H2A.Z, the three phased nucleosomes were present only after PEG treatment (Fig. 7F). Conversely, in the genes with decreased H2A.Z, the three phased nucleosome disappeared after PEG treatment. These demonstrate that H2A.Z deposition is associated with nucleosome phasing in response to PEG treatment in root. As H2A.Z deposition is related to down-regulation of PEG-responsive genes in root (Fig. 5D), we inferred that nucleosome phasing is also negatively involved in gene regulation.

***Cis*-elements enriched in the promoter of genes with changed histone marks**

Through above observations that the responses of chromatin dynamics to developmental or environmental cues were distinct in leaf and root, we considered that tissue-specific factors may direct the change of chromatin structure. Transcription factors (TFs) are the best candidates for the directing factors. *Cis*-elements analysis of promoter of genes (1.5kb upstream from TSS) with changed histone marks was performed to predict which kinds of

TFs are responsible for directing the writing or erasing of histone marks in different tissues. The five most enriched *cis*-elements for each variation of histone marks were shown in Fig. 8. For developmental genes, the *cis*-elements of Dof, AT-hook, MADS families TFs were associated with most of histone mark variations (Fig. 8A), suggesting that they were generally involved in the deposition or removal of histone marks in different tissues. The C2H2 Zinc finger TFs binding *cis*-element was inclined to be enriched in the genes with down-regulated histone marks in leaf (Fig. 8A), making it a good candidate as the tissue-specific guider for changing histone marks. In addition, ARID/BRIGHT TFs were responsible for some of histone mark variation (Fig. 8A). More specifically, YABBY TFs were associated with deposition of H3K36me3 in leaf (Fig. 8A). For PEG-induced genes, Dof, AT-hook, C2H2 Zinc finger and AP2 TFs were identified to be the most enriched *cis*-elements binding TFs (Fig. 8B). The bZIP TFs were more likely to be associated with up-regulation of H3K4me3 in leaf (Fig. 8B). MADS box *cis*-element was mostly enriched in the genes with decreased H3K9ac in leaf (Fig. 8B). Surprisingly, we found that the TATA box in many genes was associated with decrease of H3K9ac in leaf (Fig. 8B). It is not likely that TBP proteins guide histone deacetylation in these genes. Instead, we speculate that maybe the decrease of H3K9ac blocks TBP binding to TATA-containing promoters in leaf.

Discussion

The unique feature of H3K27me3 enrichment and function in sorghum

In this study we investigate the genome-wide enrichment of histone marks including H3K9ac, H3K27ac, H3K4me3, H3K36me3, H3K4me2, H3K27me3 and H2A.Z in sorghum. As reported in other plants, the distribution patterns of these marks in intragenic region are conserved in sorghum. For instance, H3K9ac, H3K27ac and H3K4me3 tend to be enriched in the TSS region, and H3K36me3, H3K4me2, H3K27me3 and H2A.Z were distributed in the whole transcribed region (Fig. 1A and 1B). However, early report revealed that H3K27me3 enrichment peaks in *Arabidopsis* was much shorter than that in animals (usually 20-50kb) and limited to single genes mostly [14]. But our analysis in *Arabidopsis*, rice and sorghum indicates that there are considerable long H3K27me3 regions in these species despite much less than the short regions (Fig. 2A and Additional file 1: Fig. S3A). The maximum length of H3K37me3 region can reach more than 200 kb in sorghum. These results suggest that H3K27me3 enrichment can also spread to a long region covering multiple genes in plants. In Drosophila, spreading of H3K27me3 requires specific *cis*-regulatory sequences, called Polycomb response elements (PREs) and pre-existing H3K27me3 [29]. In mammalian, spatially interacted nucleation sites are targeted by PRC2 complex to create H3K27me3-forming Polycomb foci for initiation of H3K27me3 spreading [30]. Indeed, we found four H3K27me3 islands where the long H3K27me3 regions were clustered on three chromosomes of sorghum but not in *Arabidopsis* and rice (Fig. 2C and Additional file 1: Fig. S3C). It is intriguing to unveil whether the H3K27me3 islands in sorghum are generated by spreading from PREs as in Drosophila or nucleation sites as in mammalian in the future. It has been disclosed that H3K27me3-mediated chromatin interaction takes place in *Arabidopsis* [31, 32], but whether the interaction of chromatin is the prerequisite for H3K27me3 spreading remains to be elucidated. In our opinion, analysis of 3D genome structure of sorghum to reveal whether the four H3K27me3 islands are interacted is necessary for the next step.

Besides, LHP1, the reader of H3K27me3 in plant, has been reported to be responsible for the spreading of H3K27me3 towards the 3' end of the gene body in *Arabidopsis* [33]. It is of great importance to investigate whether the homolog of LHP1 in sorghum is involved in the spreading of H3K27me3 enrichment to a longer region.

In general, H3K27me3 has been considered as a repressive histone mark mainly engaging in gene silencing. However, our results showed that the deficiency of long H3K27me3 regions was not always accompanied by activation of all the covering genes (Additional file 1: Fig. S9B), suggesting the loss of H3K27me3 is not sufficient for the activation of these genes. Moreover, we found differential expression of many genes where H3K27me3 did not change in root and leaf (Additional file 1: Fig. S7A-S7C). And few genes induced by PEG treatment requires the decrease of H3K27me3 (Fig. 5C and 5D). These results suggest that H3K27me3 removal is not necessary for part of gene induction. In *Arabidopsis*, H3K27me3 is not removed in *FLOWERING LOCUS T* (*FT*), the prominent florigen, during the induction of the gene [34]. Although H3K27me3 level is decreased in *FLC* after it is transcriptionally activated, the decrease of the mark is obviously later than the activation of transcription [35]. These phenomena make it necessary for us to reconsider the role of H3K27me3 in gene expression regulation. More surprisingly, recent study in mammalian showed that H3K27me3 was also involved in the promotion of gene expression. The deposition of H3K27me3 attenuates binding of repressor SP1 to the promoter of the target gene thus facilitating transcription [36]. EZH2, a H3K27me3 methylase, was revealed to be a transcription activator to activate gene expression but independent of its methyltransferase activity [37]. In plant, LHP1 was reported to be a transcriptional activator as well. However, whether the function of transcriptional activation of LHP1 is related to its binding to H3K37me3 is not clear. In this study we found many TSS-enriched H3K27me3 peak in leaf but not in root (Fig. 1B). The differential enrichment of these peaks was related to the gene up-regulation in leaf (Additional file 1: Fig. S8A). However, the possibility could not be ruled out that other active histone marks are involved in these gene induction because the enrichment of many active histone marks were also enhanced in these genes (Additional file 1: Fig. S8E). Although positive relationship between changes of H3K27me3 enrichment and gene expression was also observed in root, H3K27me3 change in these genes is not gene-specific but may be in the context of variation of long H3K27me3 regions (Additional file 1: Fig. S8C). Therefore, the relationship between increase of H3K27me3 with gene up-regulation in root is not convincing. We speculate that leaf-specific factors, which is similar to cMyc interacting with EZH2 in animal [37], may guide the deposition of H3K27me3 in the TSS region correlated with gene activation. How depletion of H3K27me3 methylases in sorghum affects expression of these genes possibly regulated by H3K27me3 positively is yet to be determined to prove the view.

H3K36me3 deposition inhibit both H3K27me3 and H2A.Z enrichment

Inhibition of PRC2 activity by H3K36m3 has been reported in animal and plant by biochemical analysis [21, 22], but whether H3K36me3 can serve as an active mark to restrict H3K27me3 spreading is not clear. Comparative analysis of genome-wide distribution of the two marks is missing. Here, by comparing genome-wide distribution of 7 histone marks in intragenic regions we found that gene-body-enriched H3K27me3 is depleted in the genes marked with H3K9ac, H3K27ac, H3K36me3 (Fig. 3A, 3B and 3D). But among these active

marks only H3K36me3 is distributed in the entire gene body region which is similar to the distribution pattern of H3K27me3 (Fig. 3D and 3G), making it a suitable marker for inhibiting H3K27me3 deposition. When we visualized H3K36me3 enrichment in H3K27me3 island, we obtained mutually exclusive distribution patterns of H3K27me3 and H3K36me3 (Fig. 3H). In the long H3K27me3 regions, a gap would be present where H3K36me3 is deposited, indicating the inhibitory effect of H3K36me3 deposition on H3K37me3 modification. But on the other hand, this also demonstrates that H3K36me3 cannot prevent the spreading of H3K27me3 since H3K27me3 is abundant at both sides of H3K36me3 region. It is of great interest to explore how H3K27me3 spreads over H3K36me3 regions. We also observed that strong enrichments of H3K36me3 were most associated with weak H2A.Z enrichments in the TSS region but hardly associated with gene-body-enriched H2A.Z (Fig. 3D and 3E). This reminds us of the antagonistic relationship between H2A.Z and DNA methylation in *Arabidopsis* [26]. We found that similar to H3K27me3, strong enrichment of H3K36me3 disrupted the long H2A.Z regions (Fig. 3I), documenting that the enrichment of H3K36me3 also inhibit H2A.Z deposition. In mammals, H3K36me2/3 serves as platform to recruit DNMT3A for DNA methylation in the gene body in order for preventing spurious transcription initiation [38], but the scenario has not been revealed in plants. It is possible that one of the main functions of H3K36me3 and DNA methylation is to promote gene body H2A.Z eviction in plants. In *Arabidopsis* enrichment of H3K27me3 and deposition of H2A.Z are mutually reinforced [39], raising the possibility that the active marks H3K36me3 and DNA methylation can counteract both H2A.Z and H3K27me3. However, our results showed that H3K36me3 variation overlapped with part of changes of H2A.Z and H3K27me3 between leaf and root (Fig. 4C), suggesting that H3K36me3 deposition is not the main cause of removal of H2A.Z and H3K27me3.

The role of histone marks in the regulation of developmental and stress-responsive genes

We compared genome-wide enrichment of histone marks in leaf and root to unravel to what extent do histone marks affect transcriptional reprogramming of developmental genes. The number of differential intragenic enrichments of each histone marks is much less than that of DEGs. Nearly half of DEGs did not concern change of any histone marks (Additional file 1: Fig. S7A), suggesting that specific genes were epigenetically regulated in response to developmental cues. In addition, we found that most of histone mark changes showed small proportions of overlapping between each other (Fig. 4C). This demonstrate that different histone marks are involved in the regulation of divergent subsets of developmental genes. The dynamic change of histone marks in response to PEG in leaf and root was also compared in this study. H3K9ac, H3K27ac, H3K4me3, H3K4me2, and H2A.Z are associated with gene induction in varying degree in response to PEG (Fig. 5A-5D). The fold changes of the genes with changed histone marks were greater than those of the genes with unchanged histone marks or without enrichment of histone marks (Additional file 1: Fig. S10), suggesting that the variation of histone marks can improve the degree of gene induction. However, H3K36me3 and H3K27me3 enrichments in few DEGs changed both in leaf and root after PEG treatment (Fig. 5C and 5D), suggesting that these two marks are probably not involved in PEG-responsive gene regulation in the process of response in sorghum. The study in moss Physcomitrella patens also showed that H3K27me3 change was not significant in response to drought stress [40]. However, in *Arabidopsis* salt priming results in shortening and

fractionation of H3K27me3 islands [41]. The H3K27me3 dynamics reponsding to salt and cold stress has also been revealed in rice [42, 43]. These demonstrate that different stress signals in different species induce distinct responses of H3K27me3 landscape.

Whether histone marks act as the cause of gene expression changes or as the consequence of variation of transcriptional activity is still controversial. Our results showed that for some of the common DEGs that are both developmentally regulated or induced by PEG, histone marks were differentially enriched between leaf and root but would not change by responding to PEG treatment (Fig. 6A and 6B). This indicates that histone marks can not be the tags of changes in the transcriptional activity of these genes, and are not necessary for the stress-responsive induction of these genes as well. One of the possibilities is that variation of histone marks in these genes is directed by development-related but not stress-related TFs. Indeed, many writers or erasers of histone marks have been reported to interact with specific TFs in plants [44-49]. By *cis*-element analysis of promoter of genes with changed histone marks we identified several different TFs that might be responsible for guiding the writing and erasing of histone marks (Fig. 8A and 8B). Future works should be focus on the confirmation of development-specific or stress-specific TFs and clarifying how they are differentially involved in dynamic change of different histone mark in response to developmental and environmental signals.

Conclusions

In this study, we revealed the existence of a great number of long H3K27me3 regions in sorghum, part of which cluster in four areas of the genome. Besides, differential enrichment of H3K27me3 occurs in both short and long regions between root and leaf. However, the decrease of long H3K27me3 regions could not give rise to the increased expression of all the covered genes. H3K36me3 play part of roles in inhibiting the deposition of H3K27me3 but seems not to be involved in preventing the spreading of the long H3K27me3 regions. It requires further investigation whether the similar mechanism is exploited for de novo modification and spreading of H3K27me3 in both short regions or single genes and long regions. In addition, we also found the genome-wide antagonistic relationship between H3K36me3 enrichment and H2A.Z deposition, which is similar as the relationship between DNA methylation and H2A.Z in *Arabidopsis*. It is of great interest to unveil how these marks coordinate or counteract with each other to regulate gene expression.

Methods

Plant growth conditions and PEG treatment

Sorghum (BTx623 variety) seeds were surface-sterilized and then soaked for germination. The germinated seeds were transferred to the nursery box (1/2 MS) to continue growing in a growth room kept at 28□ with a 12h light/12 dark cycle. The 14-day-old seedlings were transferred to the nursery box containing 20% PEG6000 that was prepared with 1/2 MS. The seedlings transferred to fresh 1/2MS solution was used as the control. After 6 hours treatment, the third leaves and roots were harvested respectively for subsequent experiments.

RNA-seq

The leaves and roots were harvested and frozen immediately in liquid nitrogen. Total

RNA was extracted using TRNzol Universal Reagent (TIANGEN) following the manufacturer's instructions. The quality of RNA was measured by gel electrophoresis, Nanodrop analyzer, Labchip and Qubit analyzer respectively. mRNA was purified by using oligo(dT) and then fragmented by incubating in fragmentation buffer. The fragmented mRNA was primed with random hexamer primers and reverse-transcribed with Reverse Transcriptase. After end repair, adenylation, adaptor ligation, purification,PCR amplification, and quality control, sequencing was performed on an Illumina HiSeq system.

ChIP-seq

Chromatin immunoprecipitation (ChIP) was performed as described in Hu et al. [3], In brief, the samples were crosslinked with 1% (v/v) formaldehyde under vacuum for 30min and then stopped by adding 2M glycine for another 5min. The nuclei were extracted using extraction buffer I (0.4M sucrose, 10mM Tris HCl pH=8.0, 10mM $MgCl_2$, 5mM β-mercaptoethanol, 0.1mM PMSF, Protease inhibitor), extraction buffer II (0.25M sucrose, 10mM Tris HCl pH=8.0, 10mM $MgCl_2$, 1% Triton X-100, 5mM β-mercaptoethanol, 0.1mM PMSF, Protease inhibitor) and extraction buffer III (1.7M sucrose, 10mM Tris HCl pH=8.0, 2mM $MgCl_2$, 0.15% Triton X-100, 5mM β-mercaptoethanol, 0.1mM PMSF, Protease inhibitor), and split in nuclei lysis buffer (50mM Tris HCl pH=8.0, 10mM EDTA, 1% SDS, Protease inhibitor). The chromatin DNA was broken into 200-500bp fragments by sonication (Soniprep 150, energy 5, 15s on and 1min off for 5 times). The antibodies H3K9ac (07-352, Millipore), H3K27ac (07-360, Millipore), H3K4me2 (07-030, Millipore), H3K4me3 (07-473, Millipore), H3K27me3 (A2363, ABclonal), H3K36me3 (ab9050, abcam), H2A.Z (preparation using antigen "GKGLLAAKTTAAKSAEKDKGKKAPV") were incubated with Protein A agarose (16-157, Millipore) in ChIP dilution buffer (1% Triton X-100, 1.2mM EDTA, 16.7mM Tris HCl pH=8.0, 167mM NaCl) before mixed with sonicated chromatin. The immunoprecipitated agarose beads were washed with low salt buffer (150mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris HCl pH=8.0), high salt buffer (500mM NaCl, 0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris HCl pH=8.0), LiCl buffer (0.25M LiCl, 1% NP-40, 1% sodium deoxycholate, 1mM EDTA, 10mM Tris HCl pH=8.0) and TE buffer (10mM Tris HCl pH=8.0, 1mM EDTA), and then eluted with elution buffer (1% SDS, 0.1M $NaHCO_3$). After de-crosslinking, The eluate was treated with Protease K and RNase H and DNA was purified by using chloroform: isoamyl alcohol (24:1). The libraries were constructed and then sequenced on an Illumina HiSeq system.

MNase-seq

The samples were crosslinked with 1% (v/v) formaldehyde and the nuclei were extracted as described in ChIP-seq experiment. The nuclei was resuspended with Micrococcal Nuclease (MNase) buffer, and then BSA and MNase (M0247, NEB) were added and then digested at 37□ for 5 min, 10 min, 15 min. The reaction was stopped by adding EGTA. RNase and 20ul 5M NaCl were used for removing RNA and de-crosslinking respectively by incubating at 65□ overnight. DNA was purified and separated by 2% agarose gel electrophoresis. Around 200bp DNA fragments generated by digestion of a proper time were recovered with gel extraction kit. The sequencing process is similar to that for ChIP-seq.

Raw sequencing reads filtering.

To obtain high quality clean reads of RNA-seq, ChIP-seq and MNase-seq, the raw sequencing reads were trimmed with Trimmomatic (version 0.32) [50]. The Trueseq3-PE adapters were removed using a maximum of two seed mismatches, a palindrome clip threshold of 30 and a

simple clip threshold of 10. The leading and tailing bases with quality below 20 or N bases were cut. Reads length shorter than 36 bp or with average quality per base in 4-base wide sliding window below than 15 were discarded.

Analyses of RNA-seq data.

Transcript abundance in leaf and root was quantified directly using pseudoalignment of high quality clean RNA-seq reads to the reference cDNA sequences and gene models from Sorghum_bicolor_NCBIv3 assembly of the variety BTx623 [51], as implemented in Kallisto (version 0.48.0) [52]. The output pseudoalignment bam files were converted to bigwig format using “bamCoverage” tool of deepTools (version 3.5.0) [53] for signal visualization on genome browser. The transcripts per million (TPM) and gene count matrices were respectively created by tximport (version 1.22.0) [54] in R, followed by reproducibility assessment of three replicates and detection of differentially expressed genes (DEGs) using DESeq2 (version 1.22) according to the instruction manual [55]. Genes with TPM value higher than 1 in all the three replicates of each sample were regarded as expressed genes, genes with |log2 (Fold change)| > 1, *P* adjust value < 0.05 and at least expressed in one sample were regarded as DEGs.

Analyses of ChIP-seq data.

The ChIP-seq high quality clean reads were aligned against the reference Sorghum_bicolor_NCBIv3 genome assembly using Bowtie2 (version 2.4.4) [56] with the parameters “bowtie2 --no-unal --threads 16 --sensitive -k 3 -q --phred33 --rg-id '"$i"_R1_"$i"_R2' --rg 'SM:"$i"_R1_"$i"_R2\tPL:Illumina\tLB:Illumina_1_8'”. Aligned reads with MAPQ < 5 were filtered using SAMtools (version 1.9) [57] with parameters “samtools view -F 1804 -q 5”. Duplicated alignments were removed using Picard (version 2.23.9) (http://broadinstitute.github.io/picard/) with parameters “picard MarkDuplicates REMOVE_DUPLICATES=true”. Before peak calling, the extsize were predicted by phantompeakqualtools (version 1.2.2) [58] with parameters “run_spp.R -filtchr="chrUn" -savp”. Narrow peaks of H3K9ac, H3K27ac and H3K4me3 ChIP-seq reads densities were identified using MACS2 (version 2.2.7.1) [59] with parameters “-f BAM -p 1e-5 --mfold 2 20 --nomodel --extsize $spp_result --to-large”. The input control were used by added parameter “-c”. For broad peaks calling including H3K4me2, H3K27me3, H3K36me3 and KH2A.Z, the “--broad” parameter was adding. The peaks from biological replicate samples were merged using the “merge” function of BEDTools (version 2.27) [60], two replicate bam files were merged by SAMTools. We then extracted the reads from the merged bam files that mapped on the merged peaks using the “intersect” function of BEDtools, and converted the output new bam files to RPKM normalized bigwig files with 20 bp bin size by “bamCoverage” provided by deepTools with parameters “--binSize 20 --normalizeUsing RPKM (Reads Per Kilobase per Million mapped reads) --smoothLength 50 --effectiveGenomeSize 708735318 --extendReads”. The metaplot and heat map profiles of epigenetic and genomic features were generated with “computeMatrix” tool and plotted with “plotProfile” tool of deepTools.

To evaluate the signal-to-noise ratio of the ChIP-seq datasets, the signal portion of tags (SPOT) and the fraction of reads in called peak regions (FRiP) value were calculated using Hotspot (version 4.1.0) (https://github.com/rthurman/hotspot) and commands in-house (https://github.com/hcph/Sorghum-omics-dataset-analysis). To assess the repeatability between biological samples, we converted the bam file to bigwig file and calculated the correlation coefficients between samples using multiBigwigSummary as well as visualized the results by plotCorrelation of deepTools.

Analyses of MNase-seq data.

The alignment and the mapped reads filtering process in MNase-seq are the same as that in ChIP-seq. Two replicate bam files were merged and converted to a bed file using “bamtobed” command of BEDTools. The genome-wide nucleosome position were detected using iNPS (version 1.2.2) [61] with the command line “python3 iNPS_V1.2.2.py -i INPUT.bed -o OUTPUT --s_p p”. To draw the distribution curves of nucleosomes, the upstream and downstream 1 kb regions from transcription start site (TSS) were divided into 20 bp bins using “computeMatrix” tool of deepTools with the parameters “computeMatrix reference-point --referencePoint TSS -b 1000 -a 1000 --binSize 20 --skipZeros --averageTypeBins mean ”. Then, the “plotProfile” tool of deepTools was used to visualize the distribution of nucleosomes.

Visualization of overall signals.

All types of signals were visualized using Circos (version 0.69) [62] with the commands available on github (https://github.com/hcph/Sorghum-omics-dataset-analysis).

Analyses of H3K27me3 regions.

The characteristics of H3K27me3 regions in sorghum, rice and *Arabidopsis* were analyzed. The H3K27me3 peaks in bed files and the signal normalized bigwig files were download from the ChIP-hub databases [63] using the sample ID of SRX7734769 and SRX7734770 for *Arabidopsis*, and SRX7426661 and SRX7426662 for rice. Peaks in the two replicates with 1 bp overlap (MG0) as well as overlap regions lower than 1,000 bp (MG1000) were merged using commands “bedtools merge -i INPUT.bed -c 4 -o collapse” and “bedtools merge -i INPUT.bed -d 1000”, respectively. The number of H3K27me3 regions with length distributed in 0-1, 1-2, 2-5, 5-10, 10-20, 20-50 and >= 50 kb were visualized using ggpubr (https://rpkgs.datanovia.com/ggpubr/) package in R. To investigate the distribution of H3K27m3 regions with different number of overlapped genes, we firstly found the overlapped genes for each H3K27me3 region using the “closet” command of BEDTools with parameters “bedtools closest -D ref -t all -mdb all”, then the number of overlapped genes was calculated and visualized in R. The genome-wide H3K27me3 signal of sorghum was visualized using pyGenomeTracks (version 3.5) [64]. The *Arabidopsis* and rice H3K27me3 regions longer than 10 kb were extracted by in-house command and the H3K27me3 signal located in these regions were visualized using pyGenomeTracks (https://github.com/hcph/Sorghum-omics-dataset-analysis).

Detection of differentially histone modified genes (DMGs).

Reads mapped to genes were extracted as described before. The gene location file was converted to SAF format file and the reads density on each genes was calculated by “featureCount” command of Subread (version 2.0.0) [65] with parameters “-p -P -B -C -T 4”. The TPM value was used to quantified the reads density on each gene, followed by DMGs detection using DeSeq2. Genes with TPM value higher than 10 in the two replicates were regarded as modified genes, and genes with |log2 (Fold change)| > 0.75, *P* adjust value < 0.05 and at least being modified in one sample were regarded as DMGs. To analysis the relationship between histone modification and gene expression, the pearson correlation coefficients for the log2FC of differentially modification level and differentially expression level were calculated in R. Long (>10 kb) H3K27me3 signals located in differentially H3K27me3 modified regions in roots and leaves were extracted with BEDtools as described before, the extracted H3K27me3 signal and the corresponding gene expression signal were visualized with Circlize (version 0.4.15) [66] in R. The commands for DMG detecting, correlation coefficients calculating and signal visualization are available (https://github.com/hcph/Sorghum-omics-dataset-analysis).

Detection of chromatin states.

To assess the combinatorial patterns of the seven histone modification marks, we distinct the

chromatin into 8 states using ChromHMM (version 1.17) following the instructions [67]. Genes located in each chromatin state were extracted by BEDTools and the overall expression level of genes in each chromatin state was visualized using ggpubr package in R.

Analyses of relationship between different histone modifications.

The signal matrix of the seven histone modification marks was built by “ComputeMatrix” tool of deepTools using the merged and RPKM normalized bigwig files as input with the parameters “omputeMatrix reference-point --referencePoint TSS -b 500 -a 3000 --skipZeros”. The signal of a single mark was used in turn as a template for clustering the signals of the remaining marks into three groups by “plotHeatmap” of deepTools with parameters “--colorMap viridis --zMin -3 --zMax 3 --whatToShow 'heatmap and colorbar' --clusterUsingSamples X --kmeans 3 plotFile Format pdf --dpi 300”. To compare the relationship of H3K27me3, H3K36me3 and H2A.Z in other species, the datasets in rice [27, 43, 68] and *Arabidopsis* [39, 69] were downloaded from ChIP-hub and analyzed in the same way as in sorghum. The signals located in long peaks (> 10 kb) were extracted and visualized as described before.The number of DMGs with two differential types of histone modifications or only differentially in one histone modifications in roots and leaves were calculated and visualized in R.

Motif enrichment.

The promoter sequences (upstream 1,500 bp) of DMGs were extracted using “getfasta” function of BEDTools. The sequences were submitted into the CentriMO (https://meme-suite.org/meme/tools/centrimo) online analysis toolkit for motif enrichment analysis. The “anywhere” mode was selected to perform local motif enrichment and the motif database was supported by CIS-BP 2.0 [70].

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The ChIP-seq, MNase-seq and RNA-seq data generated in this article were deposited in the NCBI Sequence Read Archive (SRA) (http://trace.ncbi.nlm.nih.gov/Traces/sra/) under the BioProject accession number PRJNA952350.

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by Research Initiation Fund for High-level Talents of China Three Gorges University.

Authors' contributions

The project was conceived by YH and supervised by XS. YH performed ChIP-seq, MNase-seq and RNA-seq with help from QH, GZ, XQ and ZD. CH performed the

bioinformatics analyses with help from YS and SB. The manuscript was drafted by YH and XS and revised by all the authors. ZH participate in the critical discussion. The authors read and approved the final manuscript.

Acknowledgements

We sincerely thank the computing platform of the National Key Laboratory of Crop Genetic Improvement in HZAU for providing the computational resources.

References

1. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K: **High-resolution profiling of histone methylations in the human genome.** *Cell* 2007, **129:**823-837.
2. Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA: **A chromatin landmark and transcription initiation at most promoters in human cells.** *Cell* 2007, **130:**77-88.
3. Hu Y, Lai Y, Chen X, Zhou DX, Zhao Y: **Distribution pattern of histone marks potentially determines their roles in transcription and RNA processing in rice.** *J Plant Physiol* 2020, **249:**153167.
4. Ding Y, Ndamukong I, Xu Z, Lapko H, Fromm M, Avramova Z: **ATX1-generated H3K4me3 is required for efficient elongation of transcription, not initiation, at ATX1-regulated genes.** *PLoS Genet* 2012, **8:**e1003111.
5. Gates LA, Shi J, Rohira AD, Feng Q, Zhu B, Bedford MT, Sagum CA, Jung SY, Qin J, Tsai MJ, et al: **Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation.** *J Biol Chem* 2017, **292:**14456-14472.
6. Wang H, Fan Z, Shliaha PV, Miele M, Hendrickson RC, Jiang X, Helin K: **H3K4me3 regulates RNA polymerase II promoter-proximal pause-release.** *Nature* 2023, **615:**339-348.
7. Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG: **H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation.** *Cell* 2013, **152:**1021-1036.
8. Sorenson MR, Jha DK, Ucles SA, Flood DM, Strahl BD, Stevens SW, Kress TL: **Histone H3K36 methylation regulates pre-mRNA splicing in Saccharomyces cerevisiae.** *RNA Biol* 2016, **13:**412-426.
9. Iwamori N, Tominaga K, Sato T, Riehle K, Iwamori T, Ohkawa Y, Coarfa C, Ono E, Matzuk MM: **MRG15 is required for pre-mRNA splicing and spermatogenesis.** *Proc Natl Acad Sci U S A* 2016, **113:**E5408-5415.
10. Li Y, Mukherjee I, Thum KE, Tanurdzic M, Katari MS, Obertello M, Edwards MB, McCombie WR, Martienssen RA, Coruzzi GM: **The histone methyltransferase SDG8 mediates the epigenetic modification of light and carbon responsive genes in plants.** *Genome Biol* 2015, **16:**79.
11. Liu B, Wei G, Shi J, Jin J, Shen T, Ni T, Shen WH, Yu Y, Dong A: **SET DOMAIN GROUP 708, a histone H3 lysine 36-specific methyltransferase, controls flowering time in rice (Oryza sativa).** *New Phytol* 2016, **210:**577-588.
12. Suzuki S, Kato H, Suzuki Y, Chikashige Y, Hiraoka Y, Kimura H, Nagao K, Obuse C, Takahata S, Murakami Y: **Histone H3K36 trimethylation is essential for multiple silencing mechanisms in fission yeast.** *Nucleic Acids Res* 2016, **44:**4147-4162.
13. Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V: **Genome-wide analysis of Polycomb targets in Drosophila melanogaster.** *Nat Genet* 2006, **38:**700-705.
14. Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE: **Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis.** *PLoS Biol* 2007, **5:**e129.
15. Hsu PL, Shi H, Leonen C, Kang J, Chatterjee C, Zheng N: **Structural Basis of H2B Ubiquitination-Dependent H3K4 Methylation by COMPASS.** *Mol Cell* 2019, **76:**712-723

e714.

16. Ringel AE, Cieniewicz AM, Taverna SD, Wolberger C: **Nucleosome competition reveals processive acetylation by the SAGA HAT module.** *Proc Natl Acad Sci U S A* 2015, **112:**E5461-5470.
17. Taverna SD, Ilin S, Rogers RS, Tanny JC, Lavender H, Li H, Baker L, Boyle J, Blair LP, Chait BT, et al: **Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs.** *Mol Cell* 2006, **24:**785-796.
18. Blackledge NP, Klose RJ: **The molecular principles of gene regulation by Polycomb repressive complexes.** *Nat Rev Mol Cell Biol* 2021, **22:**815-833.
19. Finogenova K, Bonnet J, Poepsel S, Schafer IB, Finkl K, Schmid K, Litz C, Strauss M, Benda C, Muller J: **Structural basis for PRC2 decoding of active histone methylation marks H3K36me2/3.** *Elife* 2020, **9**.
20. Schmitges FW, Prusty AB, Faty M, Stutzer A, Lingaraju GM, Aiwazian J, Sack R, Hess D, Li L, Zhou S, et al: **Histone methylation by PRC2 is inhibited by active chromatin marks.** *Mol Cell* 2011, **42:**330-341.
21. Yang H, Howard M, Dean C: **Antagonistic roles for H3K36me3 and H3K27me3 in the cold-induced epigenetic switch at Arabidopsis FLC.** *Curr Biol* 2014, **24:**1793-1797.
22. Yuan W, Xu M, Huang C, Liu N, Chen S, Zhu B: **H3K36 methylation antagonizes PRC2-mediated H3K27 methylation.** *J Biol Chem* 2011, **286:**7983-7989.
23. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, et al: **Mapping and analysis of chromatin state dynamics in nine human cell types.** *Nature* 2011, **473:**43-49.
24. Sequeira-Mendes J, Araguez I, Peiro R, Mendez-Giraldez R, Zhang X, Jacobsen SE, Bastolla U, Gutierrez C: **The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States.** *Plant Cell* 2014, **26:**2351-2366.
25. Roudier F, Teixeira FK, Colot V: **Chromatin indexing in Arabidopsis: an epigenomic tale of tails and more.** *Trends Genet* 2009, **25:**511-517.
26. Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S: **Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks.** *Nature* 2008, **456:**125-129.
27. Lu Y, Tan F, Zhao Y, Zhou S, Chen X, Hu Y, Zhou DX: **A Chromodomain-Helicase-DNA-Binding Factor Functions in Chromatin Modification and Gene Regulation.** *Plant Physiol* 2020, **183:**1035-1046.
28. Li G, Liu S, Wang J, He J, Huang H, Zhang Y, Xu L: **ISWI proteins participate in the genome-wide nucleosome distribution in Arabidopsis.** *Plant J* 2014, **78:**706-714.
29. Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, 3rd, Voigt P, Martin SR, Taylor WR, De Marco V, et al: **Role of the polycomb protein EED in the propagation of repressive histone marks.** *Nature* 2009, **461:**762-767.
30. Oksuz O, Narendra V, Lee CH, Descostes N, LeRoy G, Raviram R, Blumenberg L, Karch K, Rocha PP, Garcia BA, et al: **Capturing the Onset of PRC2-Mediated Repressive Domain Formation.** *Mol Cell* 2018, **70:**1149-1162 e1145.
31. Yang T, Wang D, Tian G, Sun L, Yang M, Yin X, Xiao J, Sheng Y, Zhu D, He H, Zhou Y: **Chromatin remodeling complexes regulate genome architecture in Arabidopsis.** *Plant Cell* 2022, **34:**2638-2651.

32. Huang Y, Sicar S, Ramirez-Prado JS, Manza-Mianza D, Antunez-Sanchez J, Brik-Chaouche R, Rodriguez-Granados NY, An J, Bergounioux C, Mahfouz MM, et al: **Polycomb-dependent differential chromatin compartmentalization determines gene coregulation in Arabidopsis.** *Genome Res* 2021, **31:**1230-1244.

33. Veluchamy A, Jegu T, Ariel F, Latrasse D, Mariappan KG, Kim SK, Crespi M, Hirt H, Bergounioux C, Raynaud C, Benhamed M: **LHP1 Regulates H3K27me3 Spreading and Shapes the Three-Dimensional Conformation of the Arabidopsis Genome.** *PLoS One* 2016, **11:**e0158936.

34. Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F: **cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis.** *Plant Cell* 2010, **22:**1425-1440.

35. Buzas DM, Robertson M, Finnegan EJ, Helliwell CA: **Transcription-dependence of histone H3 lysine 27 trimethylation at the Arabidopsis polycomb target gene FLC.** *Plant J* 2011, **65:**872-881.

36. Verma A, Singh A, Singh MP, Nengroo MA, Saini KK, Satrusal SR, Khan MA, Chaturvedi P, Sinha A, Meena S, et al: **EZH2-H3K27me3 mediated KRT14 upregulation promotes TNBC peritoneal metastasis.** *Nat Commun* 2022, **13:**7344.

37. Wang J, Yu X, Gong W, Liu X, Park KS, Ma A, Tsai YH, Shen Y, Onikubo T, Pi WC, et al: **EZH2 noncanonically binds cMyc and p300 through a cryptic transactivation domain to mediate gene activation and promote oncogenesis.** *Nat Cell Biol* 2022, **24:**384-399.

38. Yano S, Ishiuchi T, Abe S, Namekawa SH, Huang G, Ogawa Y, Sasaki H: **Histone H3K36me2 and H3K36me3 form a chromatin platform essential for DNMT3A-dependent DNA methylation in mouse oocytes.** *Nat Commun* 2022, **13:**4440.

39. Carter B, Bishop B, Ho KK, Huang R, Jia W, Zhang H, Pascuzzi PE, Deal RB, Ogas J: **The Chromatin Remodelers PKL and PIE1 Act in an Epigenetic Pathway That Determines H3K27me3 Homeostasis in Arabidopsis.** *Plant Cell* 2018, **30:**1337-1352.

40. Widiez T, Symeonidi A, Luo C, Lam E, Lawton M, Rensing SA: **The chromatin landscape of the moss Physcomitrella patens and its dynamics during development and drought stress.** *Plant J* 2014, **79:**67-81.

41. Sani E, Herzyk P, Perrella G, Colot V, Amtmann A: **Hyperosmotic priming of Arabidopsis seedlings establishes a long-term somatic memory accompanied by specific changes of the epigenome.** *Genome Biol* 2013, **14:**R59.

42. Dasgupta P, Prasad P, Bag SK, Chaudhuri S: **Dynamicity of histone H3K27ac and H3K27me3 modifications regulate the cold-responsive gene expression in Oryza sativa L. ssp. indica.** *Genomics* 2022, **114:**110433.

43. Zheng D, Wang L, Chen L, Pan X, Lin K, Fang Y, Wang XE, Zhang W: **Salt-Responsive Genes are Differentially Regulated at the Chromatin Levels Between Seedlings and Roots in Rice.** *Plant Cell Physiol* 2019, **60:**1790-1803.

44. Kim KC, Lai Z, Fan B, Chen Z: **Arabidopsis WRKY38 and WRKY62 transcription factors interact with histone deacetylase 19 in basal defense.** *Plant Cell* 2008, **20:**2357-2371.

45. Li S, Lin YJ, Wang P, Zhang B, Li M, Chen S, Shi R, Tunlaya-Anukit S, Liu X, Wang Z, et al: **The AREB1 Transcription Factor Influences Histone Acetylation to Regulate Drought Responses and Tolerance in Populus trichocarpa.** *Plant Cell* 2019, **31:**663-686.

46. Liu B, Liu Y, Wang B, Luo Q, Shi J, Gan J, Shen WH, Yu Y, Dong A: **The transcription factor OsSUF4 interacts with SDG725 in promoting H3K36me3 establishment.** *Nat Commun* 2019, **10:**2999.
47. Song ZT, Sun L, Lu SJ, Tian Y, Ding Y, Liu JX: **Transcription factor interaction with COMPASS-like complex regulates histone H3K4 trimethylation for specific gene expression in plants.** *Proc Natl Acad Sci U S A* 2015, **112:**2900-2905.
48. Wang TJ, Huang S, Zhang A, Guo P, Liu Y, Xu C, Cong W, Liu B, Xu ZY: **JMJ17-WRKY40 and HY5-ABI5 modules regulate the expression of ABA-responsive genes in Arabidopsis.** *New Phytol* 2021, **230:**567-584.
49. Yuan L, Song X, Zhang L, Yu Y, Liang Z, Lei Y, Ruan J, Tan B, Liu J, Li C: **The transcriptional repressors VAL1 and VAL2 recruit PRC2 for genome-wide Polycomb silencing in Arabidopsis.** *Nucleic Acids Res* 2021, **49:**98-113.
50. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data.** *Bioinformatics* 2014, **30:**2114-2120.
51. McCormick RF, Truong SK, Sreedasyam A, Jenkins J, Shu S, Sims D, Kennedy M, Amirebrahimi M, Weers BD, McKinley B, et al: **The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization.** *Plant J* 2018, **93:**338-354.
52. Bray NL, Pimentel H, Melsted P, Pachter L: **Near-optimal probabilistic RNA-seq quantification.** *Nat Biotechnol* 2016, **34:**525-527.
53. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T: **deepTools2: a next generation web server for deep-sequencing data analysis.** *Nucleic Acids Res* 2016, **44:**W160-165.
54. Soneson C, Love MI, Robinson MD: **Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences.** *F1000Res* 2015, **4:**1521.
55. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** *Genome Biol* 2014, **15:**550.
56. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nat Methods* 2012, **9:**357-359.
57. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics* 2009, **25:**2078-2079.
58. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, et al: **ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia.** *Genome Res* 2012, **22:**1813-1831.
59. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS: **Model-based analysis of ChIP-Seq (MACS).** *Genome Biol* 2008, **9:**R137.
60. Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.** *Bioinformatics* 2010, **26:**841-842.
61. Chen W, Liu Y, Zhu S, Green CD, Wei G, Han JD: **Improved nucleosome-positioning algorithm iNPS for accurate nucleosome positioning from sequencing data.** *Nat Commun* 2014, **5:**4909.
62. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA:

Circos: an information aesthetic for comparative genomics. *Genome Res* 2009, **19:**1639-1645.

63. Fu LY, Zhu T, Zhou X, Yu R, He Z, Zhang P, Wu Z, Chen M, Kaufmann K, Chen D: **ChIP-Hub provides an integrative platform for exploring plant regulome.** *Nat Commun* 2022, **13:**3413.
64. Lopez-Delisle L, Rabbani L, Wolff J, Bhardwaj V, Backofen R, Gruning B, Ramirez F, Manke T: **pyGenomeTracks: reproducible plots for multivariate genomic datasets.** *Bioinformatics* 2021, **37:**422-423.
65. Liao Y, Smyth GK, Shi W: **The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote.** *Nucleic Acids Res* 2013, **41:**e108.
66. Gu Z, Gu L, Eils R, Schlesner M, Brors B: **circlize Implements and enhances circular visualization in R.** *Bioinformatics* 2014, **30:**2811-2812.
67. Ernst J, Kellis M: **ChromHMM: automating chromatin-state discovery and characterization.** *Nat Methods* 2012, **9:**215-216.
68. Zhao L, Xie L, Zhang Q, Ouyang W, Deng L, Guan P, Ma M, Li Y, Zhang Y, Xiao Q, et al: **Integrative analysis of reference epigenomes in 20 rice varieties.** *Nat Commun* 2020, **11:**2658.
69. Xi Y, Park SR, Kim DH, Kim ED, Sung S: **Transcriptome and epigenome analyses of vernalization in Arabidopsis thaliana.** *Plant J* 2020, **103:**1490-1502.
70. Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, et al: **Determination and inference of eukaryotic transcription factor sequence specificity.** *Cell* 2014, **158:**1431-1443.

FIGURE LEGENDS

Fig. 1 Genome-wide changes of histone marks in response to PEG treatment in sorghum leaf and root. (**A**) Metaplots of 7 histone mark profiles of all genes in leaf and root before and after PEG treatment. (**B**) Heat maps of histone mark enrichment of intragenic region in leaf and root before and after PEG treatment. Black arrows indicate the TSS-enriched H3K27me3 present in leaf but not in root. Metaplots of nucleosome profiles near TSS of all genes with different expression levels (Q1-Q4) in root (**C**) and leaf (**D**). Q1 represents the lowest 25% of genes and Q4 represents the highest 25% of genes. The positions of +1 to +3 nucleosomes and the Nucleosome-free region (NFR) are indicated. Two-week-old sorghum seedlings were treated with 20% PEG6000 for 6 hours. Then root and leaf of control (without PEG treatment) and PEG-treated plants were harvested for ChIP-seq and MNase-seq.

Fig. 2 Characteristics of genome-wide H3K27me3 enrichment in sorghum. (**A**) Length distribution of H3K27me3 regions in leaf and root. (**B**) The number of genes covered by each H3K27me3 region in leaf and root. "MG1000" indicates that the adjacent H3K27me3 regions were joined when they were separated by less than 1 kb while "MG0" indicates that they were not jointed. (**C**) Genome browser screen shots of 4 H3K27me3 islands on Chromosomes 2, 5, 8. H3K27me3 islands were defined as the areas where a number of long H3K27me3 regions clustered.

Fig. 3 Analysis of overlapping enrichment of different histone marks in sorghum leaf. The enrichment of H3K9ac (**A**), H3K27ac (**B**), H3K4me3 (**C**), H3K36me3 (**D**), H2A.Z (**E**), H3K4me2 (**F**), and H3K27me3(**G**) in intragenic region were clustered into 3 categories (C1, C2 and C3) respectively and were put on the far left. The other 6 histone mark enrichment and gene expression level were shown next to the clustered marks. (**H**) Genome browser screen shots showing part of the areas in 4 H3K27me3 islands where H3K36me3 and H3K27me3 excluded from each other. (**I**) Genome browser screen shots showing some areas where H3K36me3 and H2A.Z excluded from each other.

Fig. 4 Correlation analysis of differential enrichment of histone marks and differential gene expression in leaf and root. (**A**) Metaplots of 7 histone mark profiles in differentially expressed genes (DEGs) between leaf and root. (**B**) Scatterplots of correlation between change of histone marks (log2FC-root vs leaf) and change of gene expression (log2FC-root vs leaf). Red dots represent the genes with both increased histone mark levels (log2FC > 0.75, $p < 0.05$) and increased expression levels (log2FC > 1, $p < 0.05$) and the genes with both decreased histone mark levels (log2FC < -0.75, $p < 0.05$) and decreased expression levels (log2FC < 1, $p < 0.05$) (Positive relationship). Blue dots represent the genes with increased histone mark levels and decreased expression levels and the genes with decreased histone mark levels and increased expression levels (Negative relationship). Gray dots represent the genes in which correlation between change of histone marks and change of gene expression was not significant. The number of the co-related genes was indicated. (**C**) The proportion of genes in which two histone marks changes between leaf and root. "Former unique" means the genes in which the former one of the two compared histone mark shown on the horizontal axis change but the latter one does not. "Latter unique" means the genes in which the latter histone mark change but the former one does not. "Common" means the genes in which both histone marks change. (**D**) Heat maps of H3K27me3 and H3K36me3 enrichments in all the genes with changed H3K36me3 (left, H3K36me3 increases in root; right, H3K36me3

increases in leaf). (**E**) Heat maps of H2A.Z and H3K36me3 enrichments in all the genes with changed H3K36me3 (left, H3K36me3 increases in root; right, H3K36me3 increases in leaf).

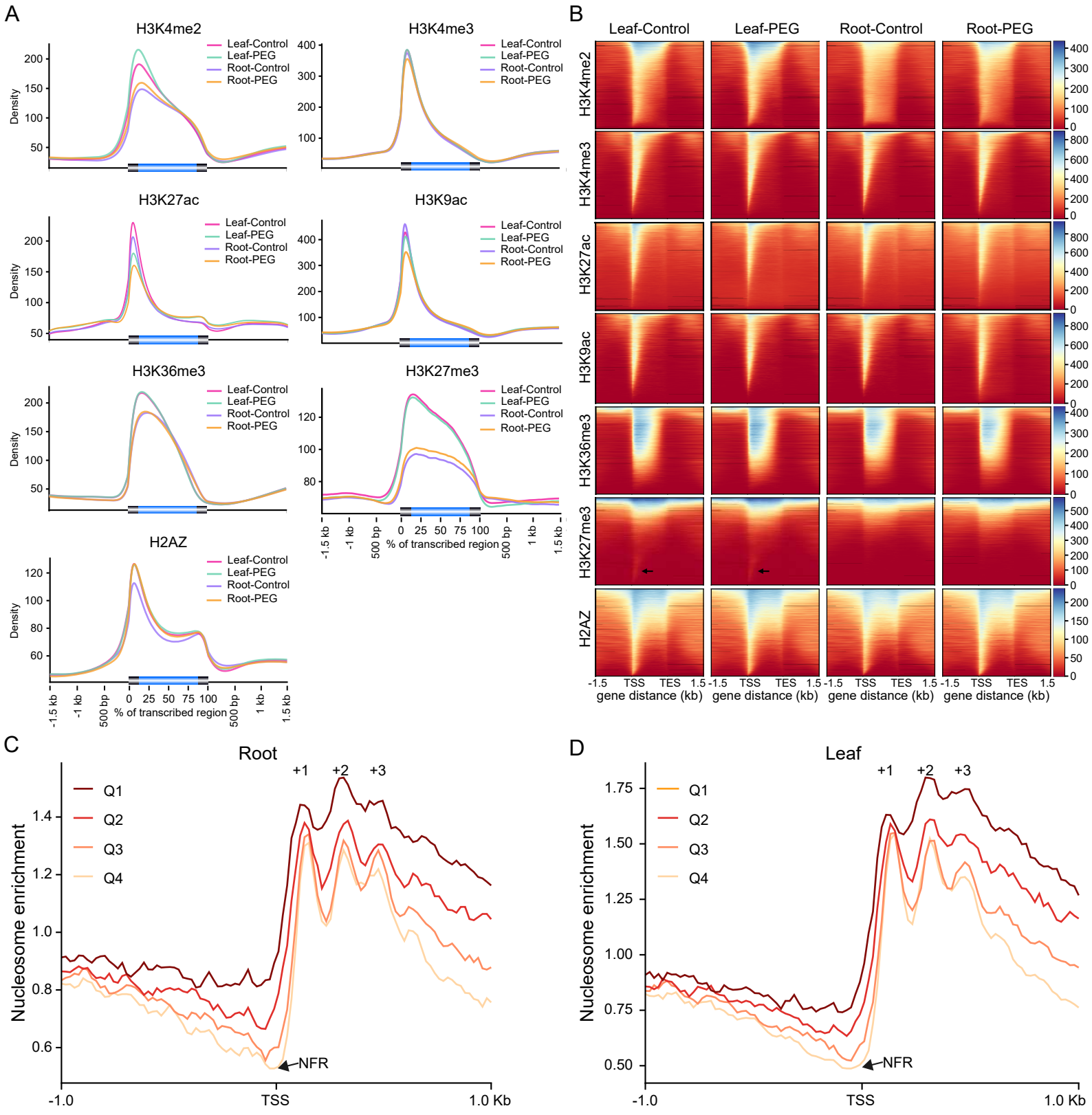
Fig. 5 Correlation analysis of changes of histone marks and gene expression in response to PEG treatment in leaf and root. Metaplots of 7 histone mark profiles in differentially expressed genes (DEGs) in response to PEG treatment in leaf (**A**) and root (**B**). Scatterplots of correlation between change of histone marks (log2FC) and change of gene expression (log2FC) in response to PEG treatment in leaf (**C**) and root (**D**). Red dots represent the genes with both increased histone mark levels (log2FC > 0.75, $p < 0.05$) and increased expression levels (log2FC > 1, $p < 0.05$) and the genes with both decreased histone mark levels (log2FC < -0.75, $p < 0.05$) and decreased expression levels (log2FC < 1, $p < 0.05$) (Positive relationship). Blue dots represent the genes with increased histone mark levels and decreased expression levels and the genes with decreased histone mark levels and increased expression levels (Negative relationship). Gray dots represent the genes in which correlation between change of histone marks and change of gene expression was not significant. The number of the co-related genes was indicated.

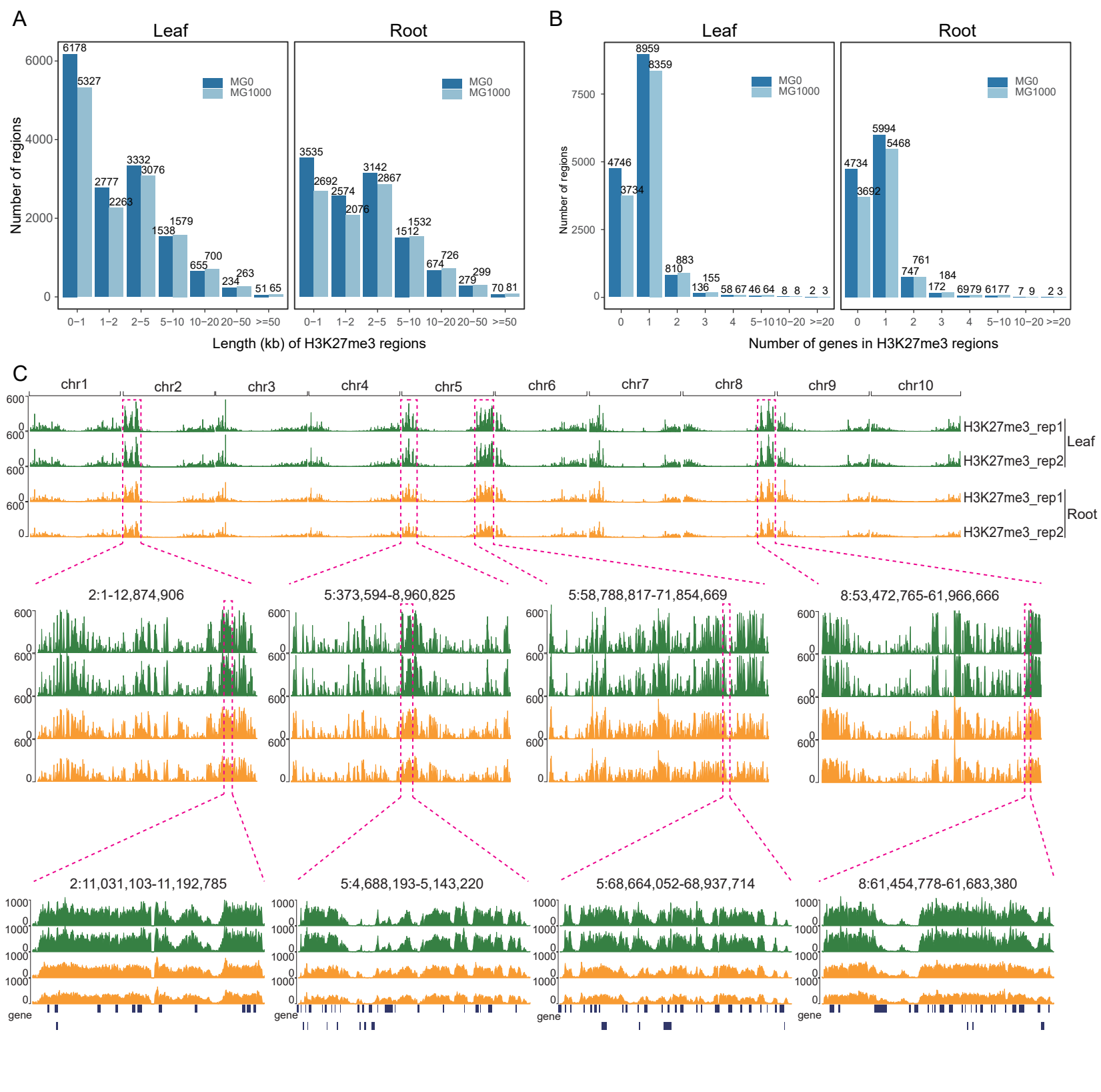
Fig. 6 Dynamic change of histone marks in the genes that were both differentially expressed between leaf and root and in response to PEG treatment. (**A**) Venn diagram of overlapping DEGs both between leaf and root and in response to PEG treatment. Four groups (I, II, III, IV) of the overlapping DEGs, also designated as the common DEGs, were identified. Root-up-DEGs: genes with increased expression in root compared to leaf; Leaf-up-DEGs: genes with increased expression in leaf compared to root; PEG-up-DEGs(leaf): genes with increased expression in response to PEG in leaf; PEG-down-DEGs(leaf): genes with decreased expression in response to PEG in leaf; PEG-up-DEGs(root): genes with increased expression in response to PEG in root; PEG-down-DEGs(root): genes with decreased expression in response to PEG in root. (**B**) The overlapping analysis of histone mark changes in four groups of the common DEGs identified in (**A**) both between leaf and root and in response to PEG treatment. “Unique” represents histone marks changed only in this condition but not the other while “common” means histone marks changed in both conditions. The number of “unique” or “common” genes was marked on the column. The total number of each group of the common DEGs was indicated.

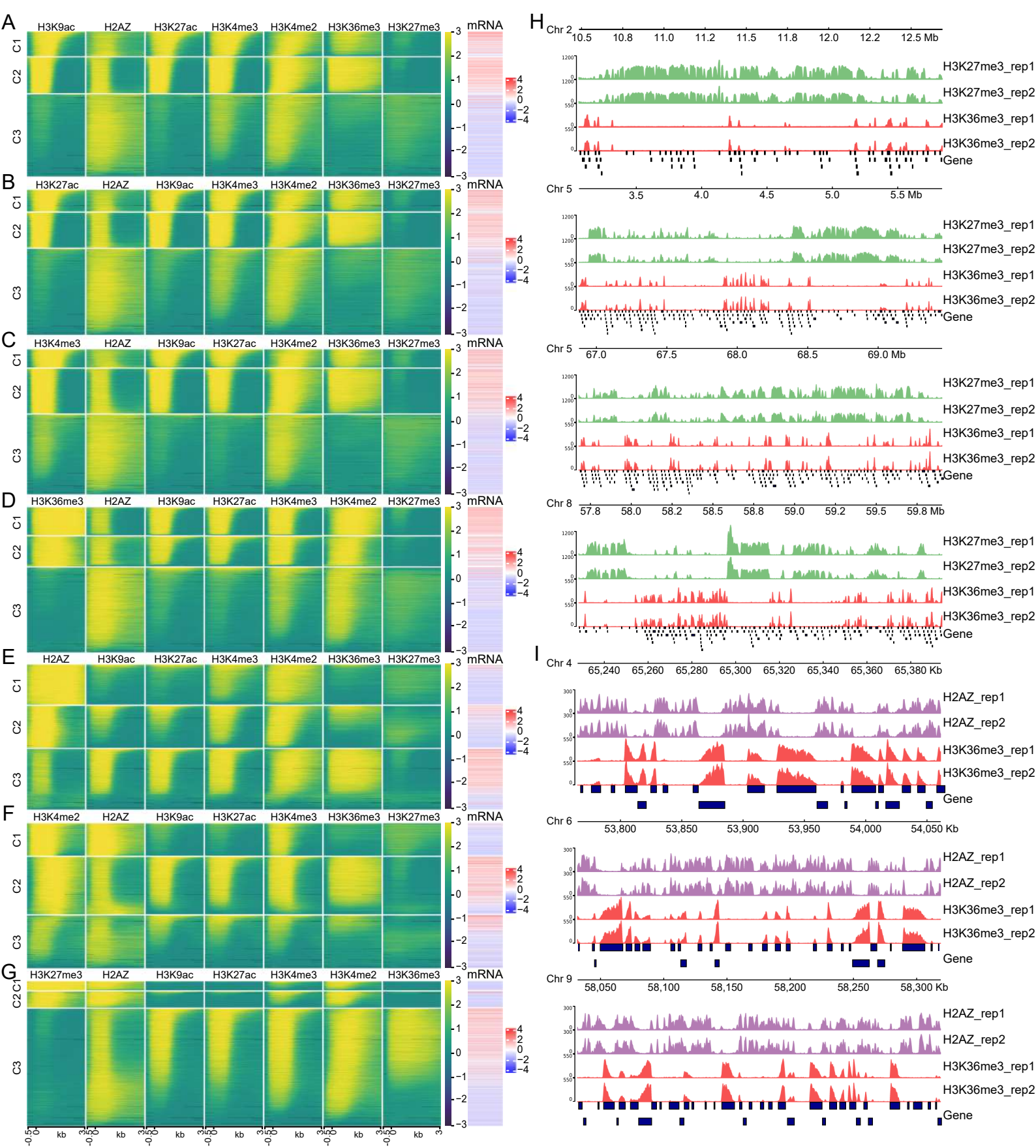
Fig. 7 Nucleosome occupancy responds to PEG treatment in leaf and root. (**A**) Metaplots of nucleosome profiles near TSS of all genes in response to PEG treatment in leaf and root. (**B**) Metaplots of nucleosome profiles near TSS of the DEGs between leaf and root. Root-up-DEGs: genes with increased expression in root; Root-down-DEGs: genes with decreased expression in root. (**C**)**,** (**D**) Metaplots of nucleosome profiles near TSS of the DEGs in response to PEG treatment in leaf (**C**) and root (**D**). PEG-up-DEGs(leaf): genes with increased expression in response to PEG in leaf; PEG-down-DEGs(leaf): genes with decreased expression in response to PEG in leaf; PEG-up-DEGs(root): genes with increased expression in response to PEG in root; PEG-down-DEGs(root): genes with decreased expression in response to PEG in root. (**E**), (**F**). Metaplots of nucleosome profiles near TSS of the genes with histone marks (H3K9ac, H3K27ac, H3K4me3 and H2A.Z) changing in response to PEG treatment in leaf (**E**) and root (**F**).

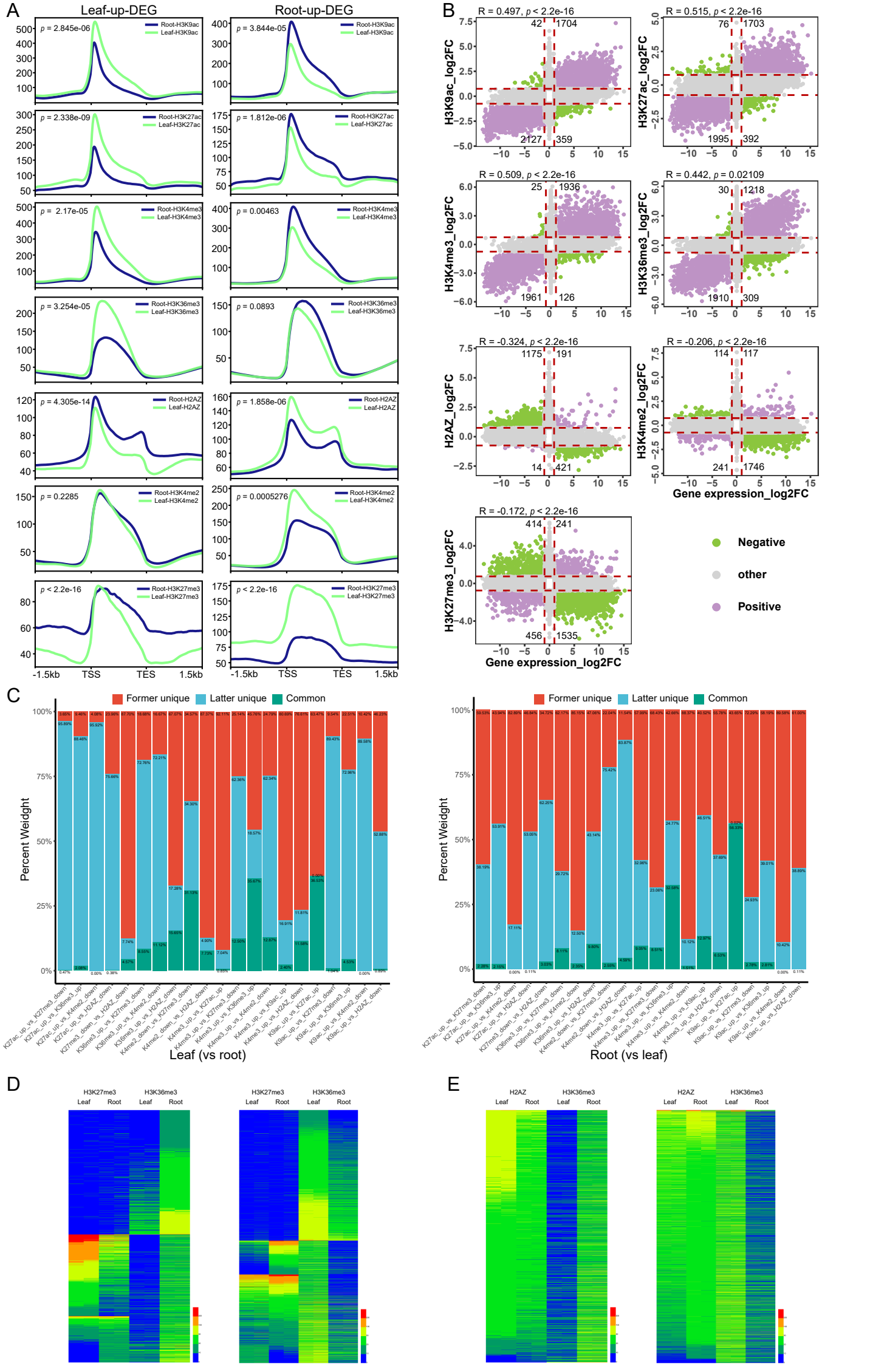
Fig. 8 *Cis*-elements analysis of genes with histone marks changing between leaf and root (A) and in response to PEG treatment (B). CL: Leaf control; PL: Leaf PEG treatment; CR:

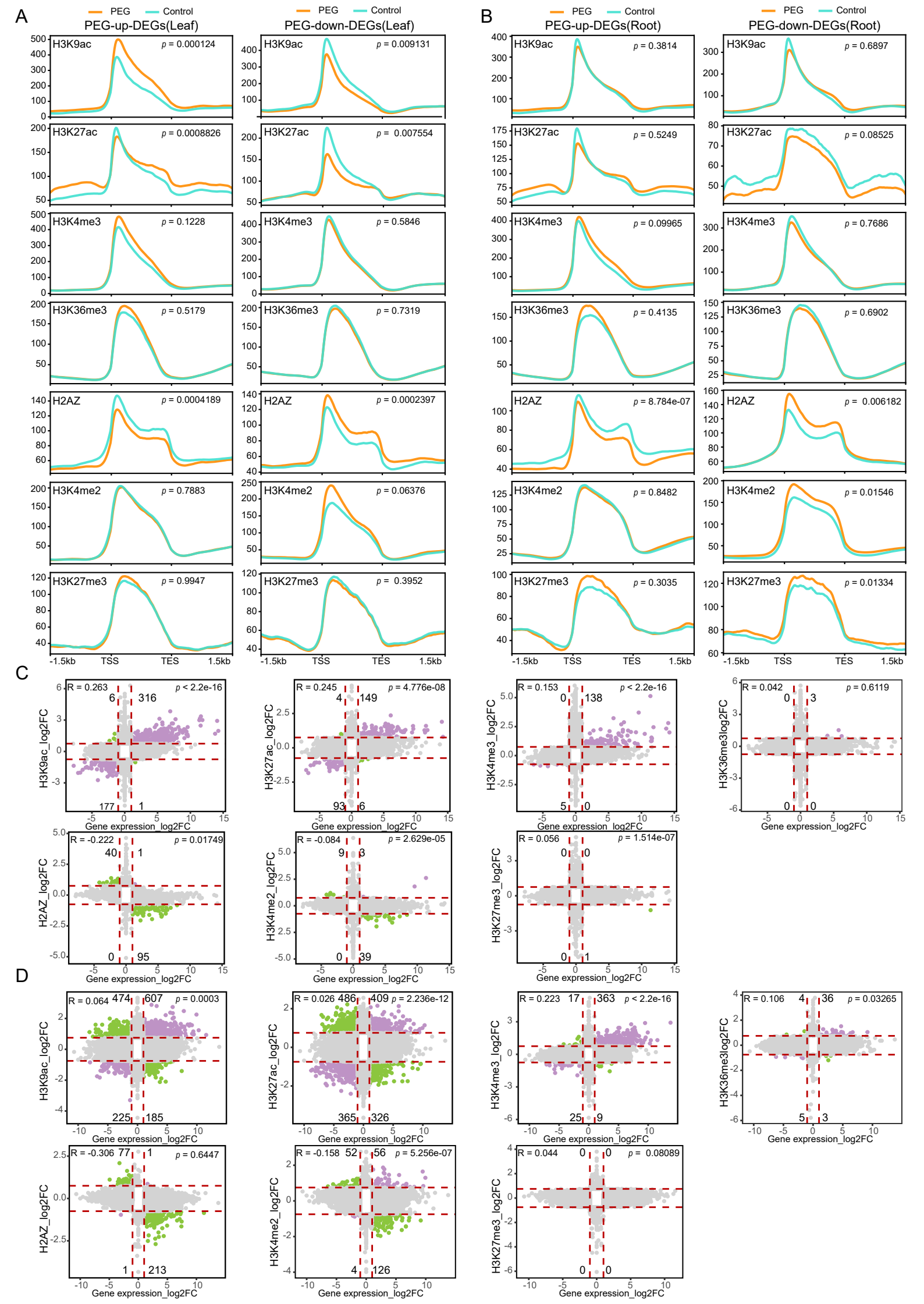
Root control; PR: Leaf PRG treatment. The percentage (%) of the genes possessing the *cis*-element relative to the total genes with the histone mark changing was marked on the column.

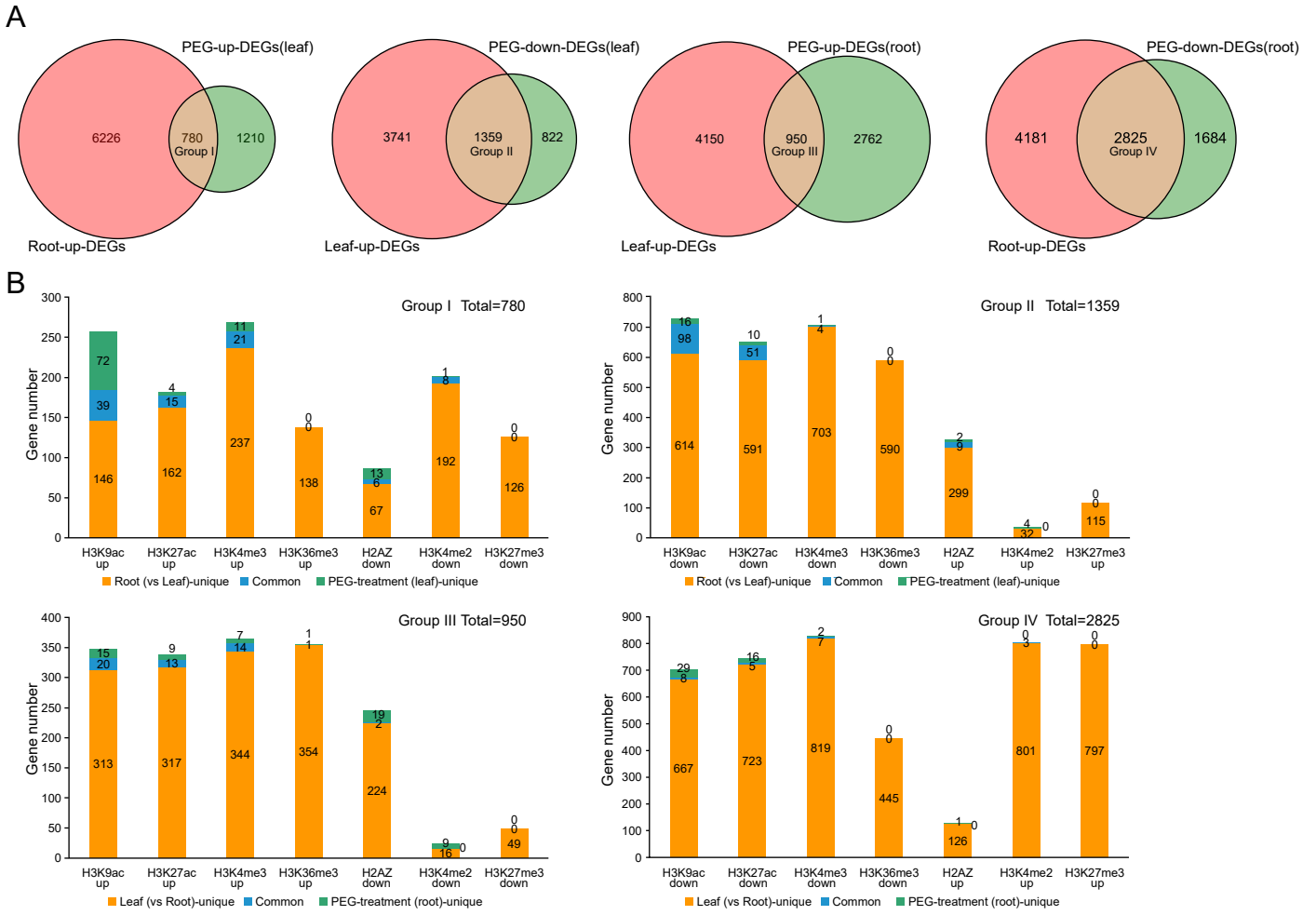

A
H3K4me2
H3K4me3
H3K27ac
H3K9ac
H3K36me3
H3K27me3
H2AZ
Leaf-Control
Leaf-PEG
Root-Control
Root-PEG
Density
% of transcribed region
-1.5 kb
-1 kb
500 bp
1 kb
1.5 kb
B
Leaf-Control
Leaf-PEG
Root-Control
Root-PEG
gene distance (kb)
TSS
TES
C
Root
Q1
Q2
Q3
Q4
+1
+2
+3
NFR
Nucleosome enrichment
-1.0
TSS
1.0 Kb
D
Leaf

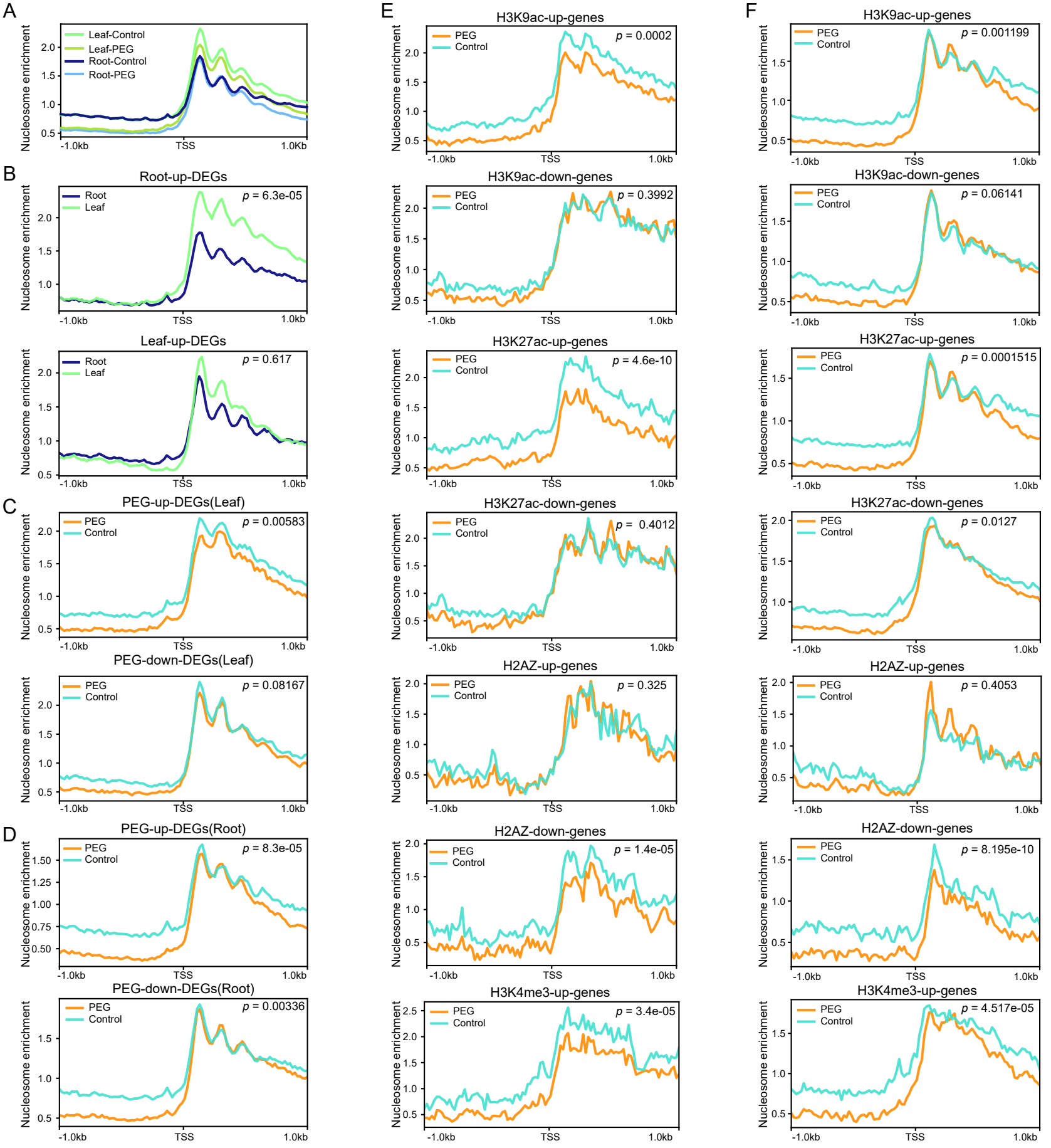

A
Leaf
Root
MG0
MG1000
Number of regions
6178
5327
2777
2263
3332
3076
1538
1579
655
700
234
263
51
65
3535
2692
2574
2076
3142
2867
1512
1532
674
726
279
299
70
81
0–1
1–2
2–5
5–10
10–20
20–50
>=50
Length (kb) of H3K27me3 regions
B
Leaf
Root
MG0
MG1000
Number of regions
4746
3734
8959
8359
810
883
136
155
58
67
46
64
8
8
2
3
4734
3692
5994
5468
747
761
172
184
69
79
61
77
7
9
2
3
0
1
2
3
4
5–10
10–20
>=20
Number of genes in H3K27me3 regions
C
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
H3K27me3_rep1
H3K27me3_rep2
Leaf
Root
2:1-12,874,906
5:373,594-8,960,825
5:58,788,817-71,854,669
8:53,472,765-61,966,666
2:11,031,103-11,192,785
5:4,688,193-5,143,220
5:68,664,052-68,937,714
8:61,454,778-61,683,380
gene


A
H3K9ac H2AZ H3K27ac H3K4me3 H3K4me2 H3K36me3 H3K27me3
mRNA
C1
C2
C3
B
H3K27ac H2AZ H3K9ac H3K4me3 H3K4me2 H3K36me3 H3K27me3
C
H3K4me3 H2AZ H3K9ac H3K27ac H3K4me2 H3K36me3 H3K27me3
D
H3K36me3 H2AZ H3K9ac H3K27ac H3K4me3 H3K4me2 H3K27me3
E
H2AZ H3K9ac H3K27ac H3K4me3 H3K4me2 H3K36me3 H3K27me3
F
H3K4me2 H2AZ H3K9ac H3K27ac H3K4me3 H3K36me3 H3K27me3
G
H3K27me3 H2AZ H3K9ac H3K27ac H3K4me3 H3K4me2 H3K36me3
kb
H
Chr 2
10.5 10.8 11.0 11.2 11.5 11.8 12.0 12.2 12.5 Mb
H3K27me3_rep1
H3K27me3_rep2
H3K36me3_rep1
H3K36me3_rep2
Gene
Chr 5
3.5 4.0 4.5 5.0 5.5 Mb
Chr 5
67.0 67.5 68.0 68.5 69.0 Mb
Chr 8
57.8 58.0 58.2 58.5 58.8 59.0 59.2 59.5 59.8 Mb
I
Chr 4
65,240 65,260 65,280 65,300 65,320 65,340 65,360 65,380 Kb
H2AZ_rep1
H2AZ_rep2
H3K36me3_rep1
H3K36me3_rep2
Gene
Chr 6
53,800 53,850 53,900 53,950 54,000 54,050 Kb
Chr 9
58,050 58,100 58,150 58,200 58,250 58,300 Kb

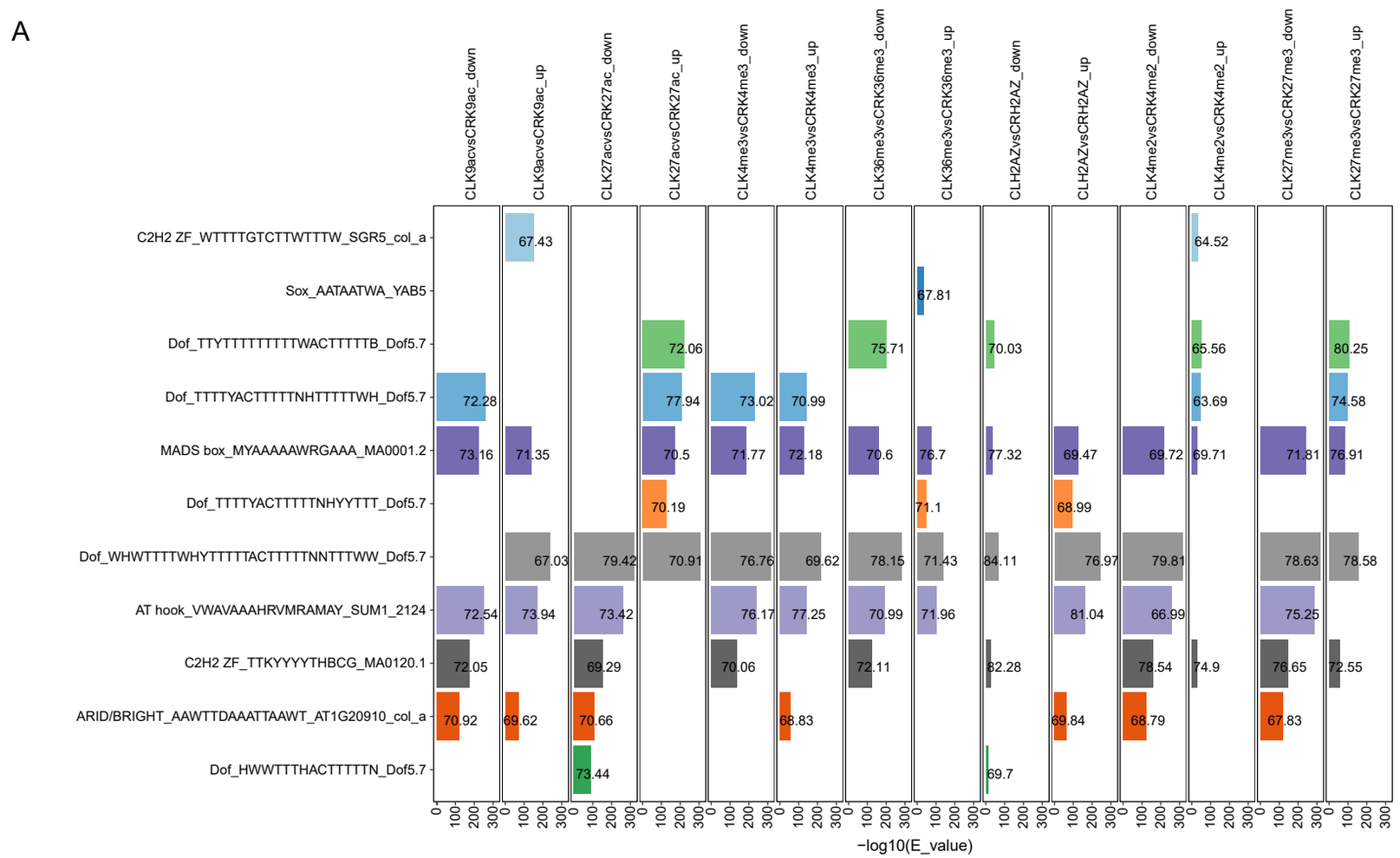

A
Leaf-up-DEG
Root-up-DEG
p = 2.845e-06
p = 3.844e-05
p = 2.338e-09
p = 1.812e-06
p = 2.17e-05
p = 0.00463
p = 3.254e-05
p = 0.0893
p = 4.305e-14
p = 1.858e-06
p = 0.2285
p = 0.0005276
p < 2.2e-16
p < 2.2e-16
Root-H3K9ac
Leaf-H3K9ac
Root-H3K27ac
Leaf-H3K27ac
Root-H3K4me3
Leaf-H3K4me3
Root-H3K36me3
Leaf-H3K36me3
Root-H2AZ
Leaf-H2AZ
Root-H3K4me2
Leaf-H3K4me2
Root-H3K27me3
Leaf-H3K27me3
-1.5kb
TSS
TES
1.5kb
B
R = 0.497, p < 2.2e-16
R = 0.515, p < 2.2e-16
R = 0.509, p < 2.2e-16
R = 0.442, p = 0.02109
R = -0.324, p < 2.2e-16
R = -0.206, p < 2.2e-16
R = -0.172, p < 2.2e-16
H3K9ac_log2FC
H3K27ac_log2FC
H3K4me3_log2FC
H3K36me3_log2FC
H2AZ_log2FC
H3K4me2_log2FC
H3K27me3_log2FC
Gene expression_log2FC
Negative
other
Positive
C
Former unique
Latter unique
Common
Percent Weidght
Leaf (vs root)
Root (vs leaf)
D
H3K27me3
H3K36me3
Leaf
Root
E
H2AZ
H3K36me3
Leaf
Root


A
PEG
Control
PEG-up-DEGs(Leaf)
PEG-down-DEGs(Leaf)
B
PEG-up-DEGs(Root)
PEG-down-DEGs(Root)
H3K9ac
H3K27ac
H3K4me3
H3K36me3
H2AZ
H3K4me2
H3K27me3
p = 0.000124
p = 0.009131
p = 0.0008826
p = 0.007554
p = 0.1228
p = 0.5846
p = 0.5179
p = 0.7319
p = 0.0004189
p = 0.0002397
p = 0.7883
p = 0.06376
p = 0.9947
p = 0.3952
p = 0.3814
p = 0.6897
p = 0.5249
p = 0.08525
p = 0.09965
p = 0.7686
p = 0.4135
p = 0.6902
p = 8.784e-07
p = 0.006182
p = 0.8482
p = 0.01546
p = 0.3035
p = 0.01334
-1.5kb
TSS
TES
1.5kb
C
R = 0.263
p < 2.2e-16
6
316
177
1
R = 0.245
p = 4.776e-08
4
149
93
6
R = 0.153
p < 2.2e-16
0
138
5
0
R = 0.042
p = 0.6119
0
3
0
0
R = -0.222
p = 0.01749
40
1
0
95
R = -0.084
p = 2.629e-05
9
3
0
39
R = 0.056
p = 1.514e-07
0
0
0
1
H3K9ac_log2FC
H3K27ac_log2FC
H3K4me3_log2FC
H3K36me3log2FC
H2AZ_log2FC
H3K4me2_log2FC
H3K27me3_log2FC
Gene expression_log2FC
D
R = 0.064
474
607
p = 0.0003
225
185
R = 0.026
486
409
p = 2.236e-12
365
326
R = 0.223
17
363
p < 2.2e-16
25
9
R = 0.106
4
36
p = 0.03265
5
3
R = -0.306
77
1
p = 0.6447
1
213
R = -0.158
52
56
p = 5.256e-07
4
126
R = 0.044
0
0
p = 0.08089
0
0

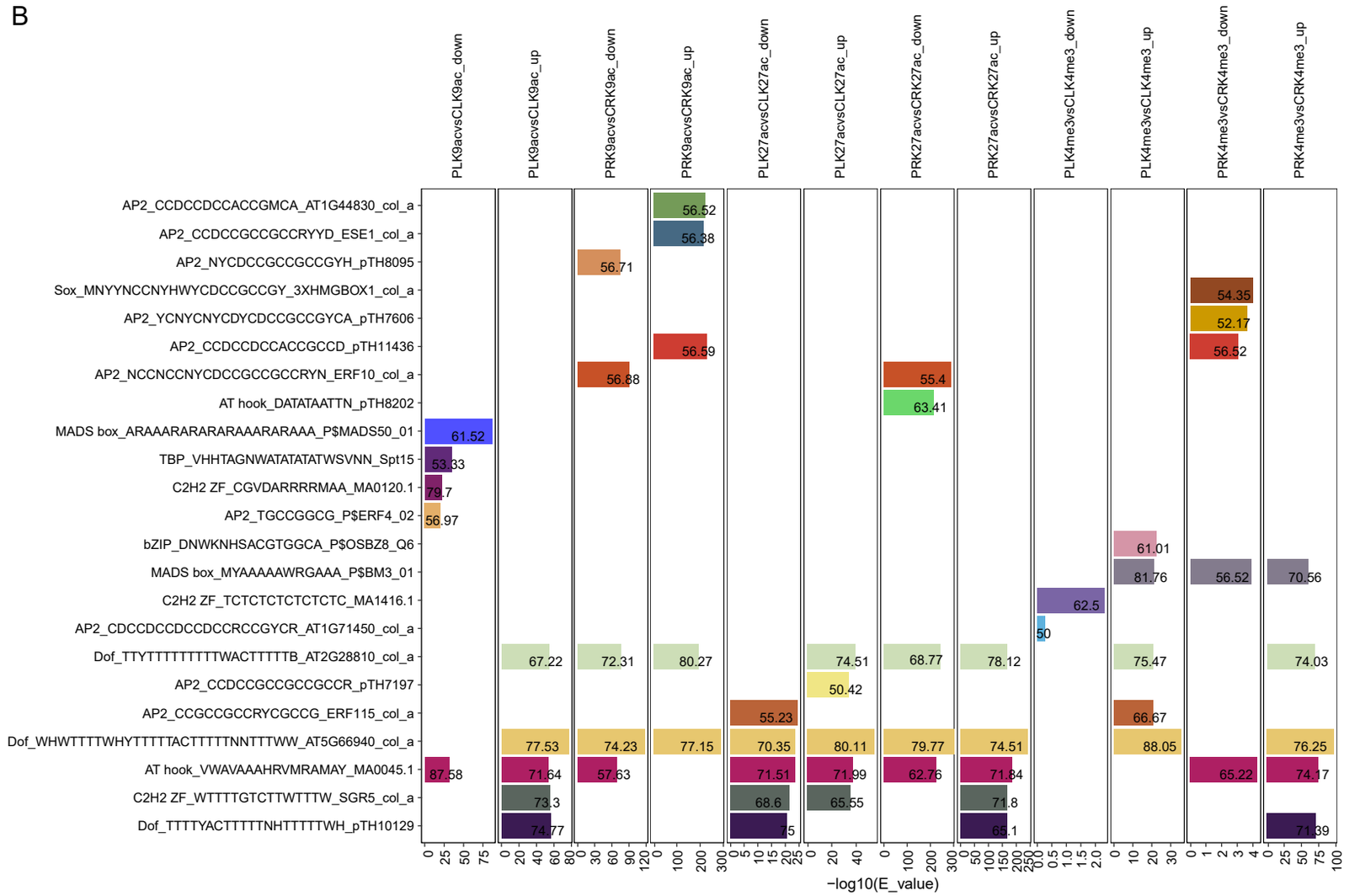

A
PEG-up-DEGs(leaf)
6226
780
Group I
1210
Root-up-DEGs
PEG-down-DEGs(leaf)
3741
1359
Group II
822
Leaf-up-DEGs
PEG-up-DEGs(root)
4150
950
Group III
2762
Leaf-up-DEGs
PEG-down-DEGs(root)
4181
2825
Group IV
1684
Root-up-DEGs
B
Group I Total=780
Gene number
H3K9ac up
H3K27ac up
H3K4me3 up
H3K36me3 up
H2AZ down
H3K4me2 down
H3K27me3 down
Root (vs Leaf)-unique
Common
PEG-treatment (leaf)-unique
Group II Total=1359
Gene number
H3K9ac down
H3K27ac down
H3K4me3 down
H3K36me3 down
H2AZ up
H3K4me2 up
H3K27me3 up
Root (vs Leaf)-unique
Common
PEG-treatment (leaf)-unique
Group III Total=950
Gene number
H3K9ac up
H3K27ac up
H3K4me3 up
H3K36me3 up
H2AZ down
H3K4me2 down
H3K27me3 down
Leaf (vs Root)-unique
Common
PEG-treatment (root)-unique
Group IV Total=2825
Gene number
H3K9ac down
H3K27ac down
H3K4me3 down
H3K36me3 down
H2AZ up
H3K4me2 up
H3K27me3 up
Leaf (vs Root)-unique
Common
PEG-treatment (root)-unique


A
Leaf-Control
Leaf-PEG
Root-Control
Root-PEG
B
Root-up-DEGs
Root
Leaf
p = 6.3e-05
Leaf-up-DEGs
Root
Leaf
p = 0.617
C
PEG-up-DEGs(Leaf)
PEG
Control
p = 0.00583
PEG-down-DEGs(Leaf)
PEG
Control
p = 0.08167
D
PEG-up-DEGs(Root)
PEG
Control
p = 8.3e-05
PEG-down-DEGs(Root)
PEG
Control
p = 0.00336
E
H3K9ac-up-genes
PEG
Control
p = 0.0002
H3K9ac-down-genes
PEG
Control
p = 0.3992
H3K27ac-up-genes
PEG
Control
p = 4.6e-10
H3K27ac-down-genes
PEG
Control
p = 0.4012
H2AZ-up-genes
PEG
Control
p = 0.325
H2AZ-down-genes
PEG
Control
p = 1.4e-05
H3K4me3-up-genes
PEG
Control
p = 3.4e-05
F
H3K9ac-up-genes
PEG
Control
p = 0.001199
H3K9ac-down-genes
PEG
Control
p = 0.06141
H3K27ac-up-genes
PEG
Control
p = 0.0001515
H3K27ac-down-genes
PEG
Control
p = 0.0127
H2AZ-up-genes
PEG
Control
p = 0.4053
H2AZ-down-genes
PEG
Control
p = 8.195e-10
H3K4me3-up-genes
PEG
Control
p = 4.517e-05
Nucleosome enrichment
-1.0kb
TSS
1.0Kb
1.0kb

A


CLK9acvsCRK9ac_down
CLK9acvsCRK9ac_up
CLK27acvsCRK27ac_down
CLK27acvsCRK27ac_up
CLK4me3vsCRK4me3_down
CLK4me3vsCRK4me3_up
CLK36me3vsCRK36me3_down
CLK36me3vsCRK36me3_up
CLH2AZvsCRH2AZ_down
CLH2AZvsCRH2AZ_up
CLK4me2vsCRK4me2_down
CLK4me2vsCRK4me2_up
CLK27me3vsCRK27me3_down
CLK27me3vsCRK27me3_up
C2H2 ZF_WTTTTGTCTTWTTTW_SGR5_col_a
Sox_AATAATWA_YAB5
Dof_TTYTTTTTTTTTWACTTTTTB_Dof5.7
Dof_TTTTYACTTTTTNHTTTTTWH_Dof5.7
MADS box_MYAAAAAWRGAAA_MA0001.2
Dof_TTTTYACTTTTTNHYYTTT_Dof5.7
Dof_WHWTTTTWHYTTTTTACTTTTTNNTTTWW_Dof5.7
AT hook_VWAVAAAHRVMRAMAY_SUM1_2124
C2H2 ZF_TTKYYYYTHBCG_MA0120.1
ARID/BRIGHT_AAWTTDAAATTAAWT_AT1G20910_col_a
Dof_HWWTTTHACTTTTTN_Dof5.7
−log10(E_value)

B


PLK9acvsCLK9ac_down
PLK9acvsCLK9ac_up
PRK9acvsCRK9ac_down
PRK9acvsCRK9ac_up
PLK27acvsCLK27ac_down
PLK27acvsCLK27ac_up
PRK27acvsCRK27ac_down
PRK27acvsCRK27ac_up
PLK4me3vsCLK4me3_down
PLK4me3vsCLK4me3_up
PRK4me3vsCRK4me3_down
PRK4me3vsCRK4me3_up
AP2_CCDCCDCCACCGMCA_AT1G44830_col_a
AP2_CCDCCGCCGCCRYYD_ESE1_col_a
AP2_NYCDCCGCCGCCGYH_pTH8095
Sox_MNYYNCCNYHWYCDCCGCCGY_3XHMGBOX1_col_a
AP2_YCNYCNYCDYCDCCGCCGYCA_pTH7606
AP2_CCDCCDCCACCGCCD_pTH11436
AP2_NCCNCCNYCDCCGCCGCCRYN_ERF10_col_a
AT hook_DATATAATTN_pTH8202
MADS box_ARAAARARARARAAARARAAA_P$MADS50_01
TBP_VHHTAGNWATATATATWSVNN_Spt15
C2H2 ZF_CGVDARRRRMAA_MA0120.1
AP2_TGCCGGCG_P$ERF4_02
bZIP_DNWKNHSACGTGGCA_P$OSBZ8_Q6
MADS box_MYAAAAAWRGAAA_P$BM3_01
C2H2 ZF_TCTCTCTCTCTCTC_MA1416.1
AP2_CDCCDCCDCCDCCRCCGYCR_AT1G71450_col_a
Dof_TTYTTTTTTTTTWACTTTTTB_AT2G28810_col_a
AP2_CCDCCGCCGCCGCCR_pTH7197
AP2_CCGCCGCCRYCGCCG_ERF115_col_a
Dof_WHWTTTTWHYTTTTTACTTTTTNNTTTWW_AT5G66940_col_a
AT hook_VWAVAAAHRVMRAMAY_MA0045.1
C2H2 ZF_WTTTTGTCTTWTTTW_SGR5_col_a
Dof_TTTTYACTTTTTNHTTTTTWH_pTH10129
−log10(E_value)