Supplementary information, Data S1

Materials and Methods

Mutant screen and map-based cloning

The *hdp1-1*, *hdp1-2* and *hdp2-1* mutants were obtained by EMS screen as described in our previous report [1, 2]. To clone the *HDP1* and *HDP2* genes, mutants were crossed with *Landsberg erecta* and hybrid F1 plants were self-pollinated to obtain the F2 population. Seedlings with long-root phenotype on 1% sucrose-containing MS media were selected from F2 plants for calculation of mutant linkage. Genomic DNA from mutants was re-sequenced to determine the location of the mutation in the mapping region.

RNA-seq analysis

For RNA-seq analysis, total RNAs were extracted from 2-week-old seedlings and submitted to RNA sequencing. Clean reads were mapped to *Arabidopsis* reference genome using TopHat. FPKM were calculated using cufflinks [3]. Pearson's coefficient plot was generated using chart. Correlation() function from the PerformanceAnalytics package. For the plot, we require: 1) lengths of genes should be longer than 500 bp; 2) FPKM of at least one sample should be higher than [3].

BiFC and subcellular localization analysis

For the BiFC assay, full-length *HDP1* and *HDP2* sequences were cloned into pSITE-nEYFP-C1 and pSITE-cEYFP-C1 vectors [4] to generate *HDP1* and *HDP2*

fused split YFP constructs. *Agrobacteria* bearing split YFP constructs were infiltrated into *N. benthamiana* leaves. YFP fluorescence was examined at 2 day-post-infiltration. For subcellular localization of HDP1 and HDP2 proteins, full-length coding sequences were cloned into pEarleyGate vectors to generate HDP1-YFP, HDP1-CFP and HDP2-YFP constructs. NbFIB2-RFP was published by Duan *et al*. [5]. *Agrobacteria* bearing these constructs were infiltrated into *N. benthamiana* leaves or transfected into *Arabidopsis* protoplasts. Fluorescence was examined at 2 day-post-infiltration in tobacco leaves or after 24 h incubation in protoplasts. For HDP1 and HDP2 subnuclear co-localization assay, nuclei were extracted from seedlings of 3FH-HDP2/HDP1-4myc hybrid plants and immunostaining was performed as reported previously [6] using anti-Flag (Sigma-Aldrich, F1804) and anti-myc (Sigma-Aldrich, C3956) antibodies.

EMSA assay

EMSA reaction was performed as described previously [7]. Full-length wild-type and mutated HDP2 coding sequences were cloned into pMAL C2X for expression of MBP fusion proteins in BL21 *E. coli* and purification as described previously [2]. DNA oligonucleotides were annealed first to generate double-stranded DNA. Then double-stranded DNA was labeled by $^{[\gamma\text{-}32]}$P-ATP using T4 polynucleotide kinase (NEB) to generate probes for EMSA reaction.

ChIP assay and ChIP-seq analysis

ChIP assay was performed as described [8] using 2-week-old seedlings. Dynabeads (Invitrogen, 10004D) were used for antibody binding. The antibodies used include, anti-pol II (Abcam, ab817), anti-H3K9me2 (Abcam, ab1220), anti-H3K4me3 (Abcam, ab8580), anti-FLAG (Sigma-Aldrich, F1804) and anti-H3K18AC (Abcam, ab1191).

For ChIP-seq analysis, Arabidopsis genome was divided into 2 kb bins. ChIP signal were calculated as follows [9, 10]:

Enrichment = $\log_2(8 + n1) – \log_2(8 + n2 * N1/N2)$

We used N1 and N2 to represent the number of aligned ChIP and input reads in each windows. N1 and N2 were the total reads in ChIP and input data, respectively. To overcome sampling noise, eight pseudo counts were added. The DMCs were defined as cytosines whose methylation level in mutant is 10% higher than that in WT plants.

Real-time qRT-PCR

For real-time qRT-PCR, total RNAs were extracted from 2-week-old seedlings using the RNeasy Plant Minikit (QIAGEN). After TURBO DNase I treatment (Ambion), 2 μg of RNA was subjected to reverse transcription reaction using the SuperScript III First-Strand Kit according to the manufacturer’s instructions (Invitrogen). The 1st-strand cDNAs were then amplified using IQ SYBR green supermix (BIO-RAD) with the CFX96 real-time PCR detection system (BIO-RAD).

Supplementary references

1 Lei M, La H, Lu K *et al.* Arabidopsis EDM2 promotes IBM1 distal polyadenylation and regulates genome DNA methylation patterns. *Proceedings of the National Academy of Sciences of the United States of America* 2014; **111**:527-532.

2 Wang X, Duan CG, Tang K *et al.* RNA-binding protein regulates plant DNA methylation by controlling mRNA processing at the intronic heterochromatin-containing gene IBM1. *Proceedings of the National Academy of Sciences of the United States of America* 2013; **110**:15467-15472.

3 Trapnell C, Roberts A, Goff L *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols* 2012; **7**:562-578.

4 Martin K, Kopperud K, Chakrabarty R, Banerjee R, Brooks R, Goodin MM. Transient expression in Nicotiana benthamiana fluorescent marker lines provides enhanced definition of protein localization, movement and interactions in planta. *The Plant journal: for cell and molecular biology* 2009; **59**:150-162.

5 Duan CG, Fang YY, Zhou BJ *et al.* Suppression of Arabidopsis ARGONAUTE1-mediated slicing, transgene-induced RNA silencing, and DNA methylation by distinct domains of the Cucumber mosaic virus 2b protein. *The Plant cell* 2012; **24**:259-274.

6 Duan CG, Zhang H, Tang K *et al.* Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. *EMBO J* 2015; **34**:581-592.

7 Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 1997; **91**:501-510.

8 Saleh A, Alvarez-Venegas R, Avramova Z. An efficient chromatin immunoprecipitation (ChIP) protocol for studying histone modifications in Arabidopsis plants. *Nature protocols* 2008; **3**:1018-1025.

9 Lang Z, Lei M, Wang X *et al.* The methyl-CpG-binding protein MBD7 facilitates active DNA demethylation to limit DNA hyper-methylation and transcriptional gene silencing. *Mol Cell* 2015; **57**:971-983.

10 Baubec T, Ivanek R, Lienert F, Schubeler D. Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell* 2013; **153**:480-492.

11 Sinzelle L, Kapitonov VV, Grzela DP *et al.* Transposition of a reconstructed Harbinger element in human cells and functional homology with two transposon-derived cellular genes. *Proceedings of the National Academy of Sciences of the United States of America* 2008; **105**:4715-4720.