Supporting Information

Ausin et al. 10.1073/pnas.1618019113

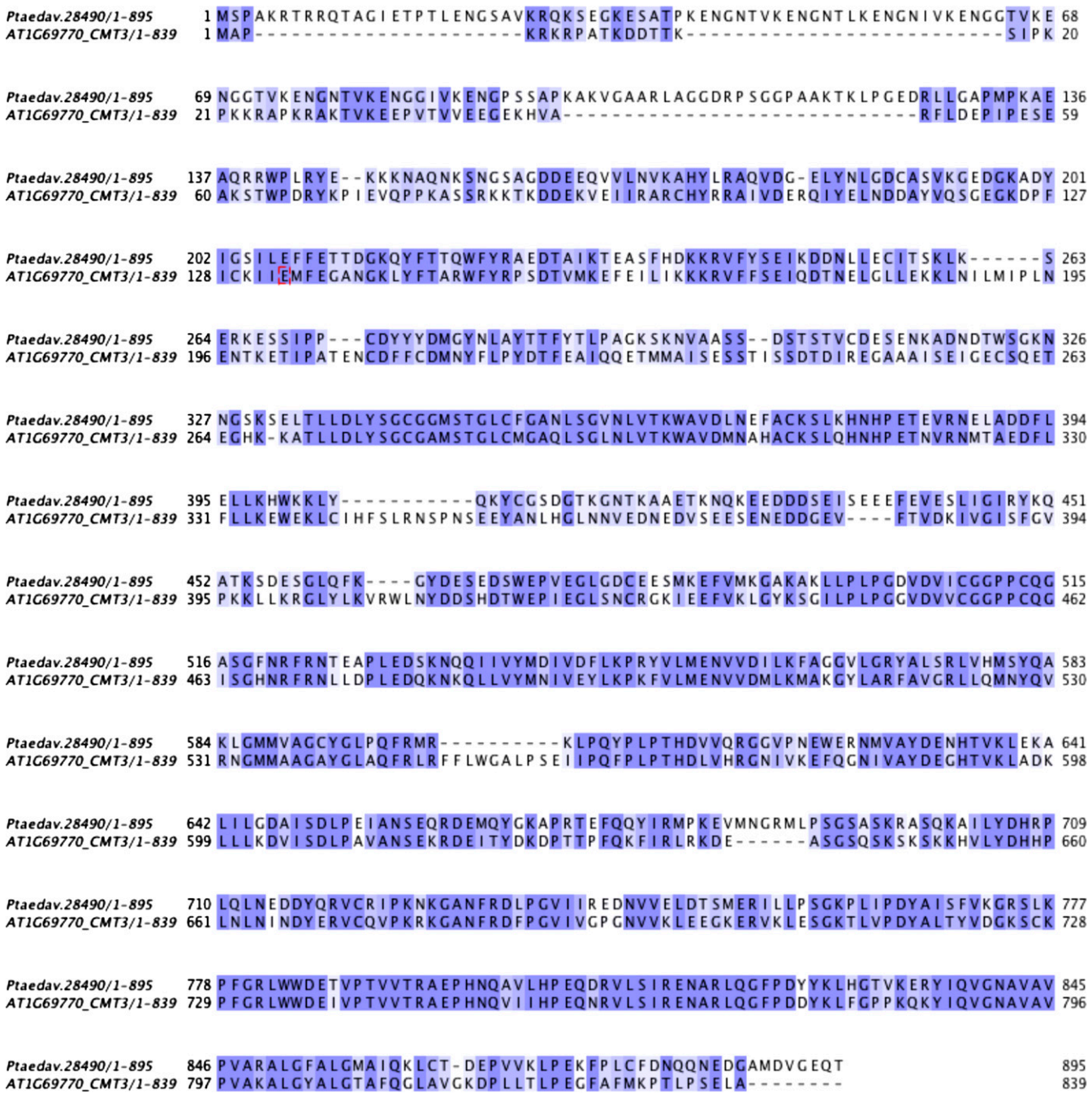

Fig. S1. Sequence alignment of *Arabidopsis* CMT3 and a putative CMT3 ortholog of *P. taedav*. Both protein sequences were extracted by reciprocal best hit of BLASTP, and then were aligned by Muscle.

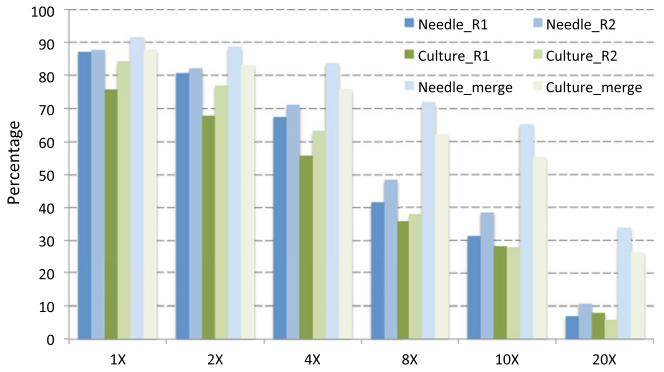

Fig. S2. BS-seq coverage shown as the percentage of cytosines that are covered by at least "X" reads. Both replicates of needle and SE culture cell were calculated separately first and then merged together as Needle/Culture_merge. Approximately 83% and 76% of total cytosines were covered by at least four reads in needle and SE culture cell, respectively.

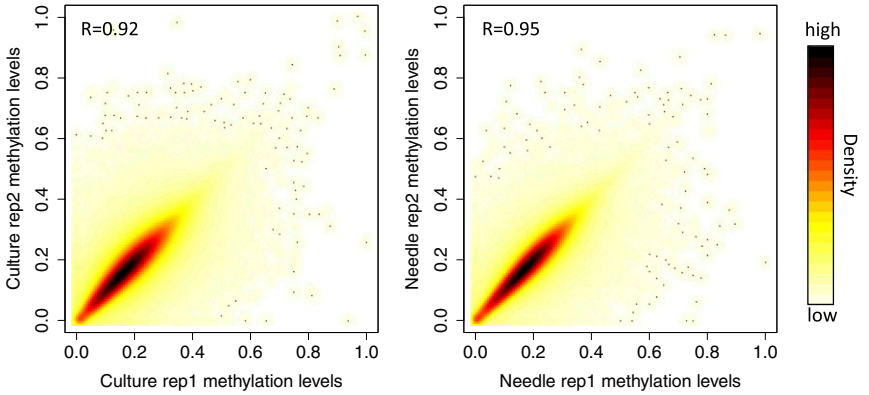

Fig. S3. Correlation between replicates of needles (*Right*) and SE culture (*Left*). Methylation levels were calculated for each 2-kbp window, and Pearson correlation coefficient was calculated and shown as R value.

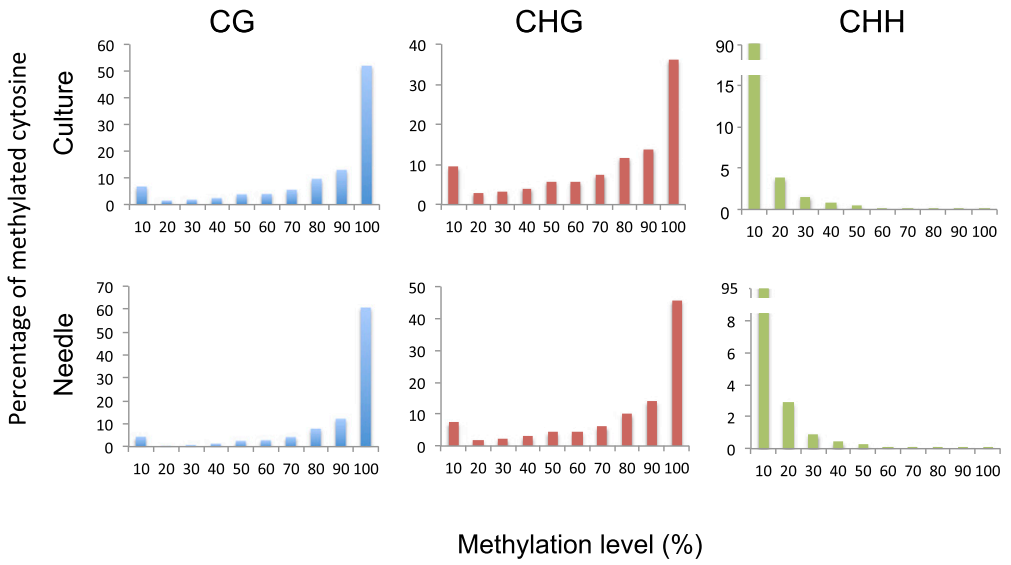

Fig. S4. Distribution of methylation levels of CG (*Left*), CHG (*Center*), and CHH (*Right*) in SE culture and needles. The methylation level is divided into 10 bins from 0 to 100%.

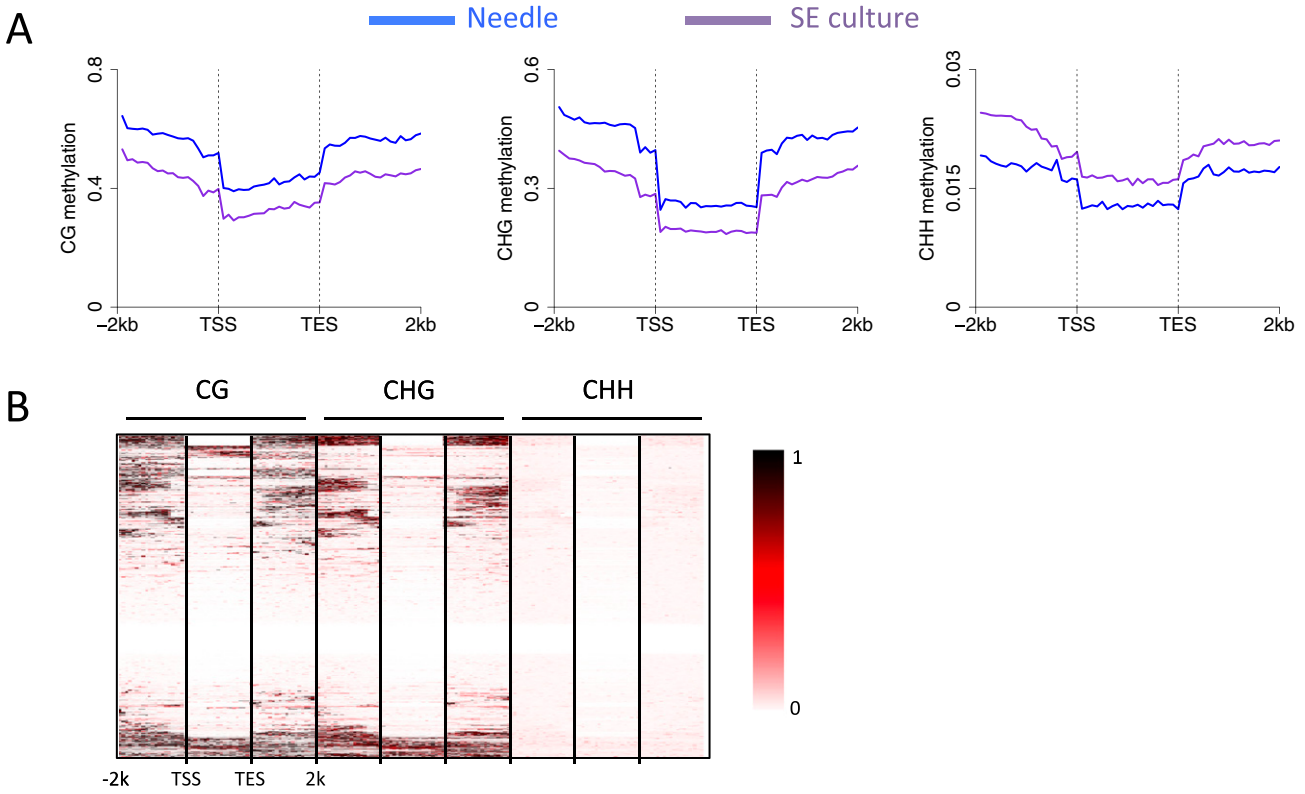

Fig. S5. Methylation patterns of genes excluding intronic TE insertions. (*A*) Metaplot of methylation in genes excluding intronic TE insertions. (*B*) Heatmap of methylation in genes excluding intronic TE insertions.

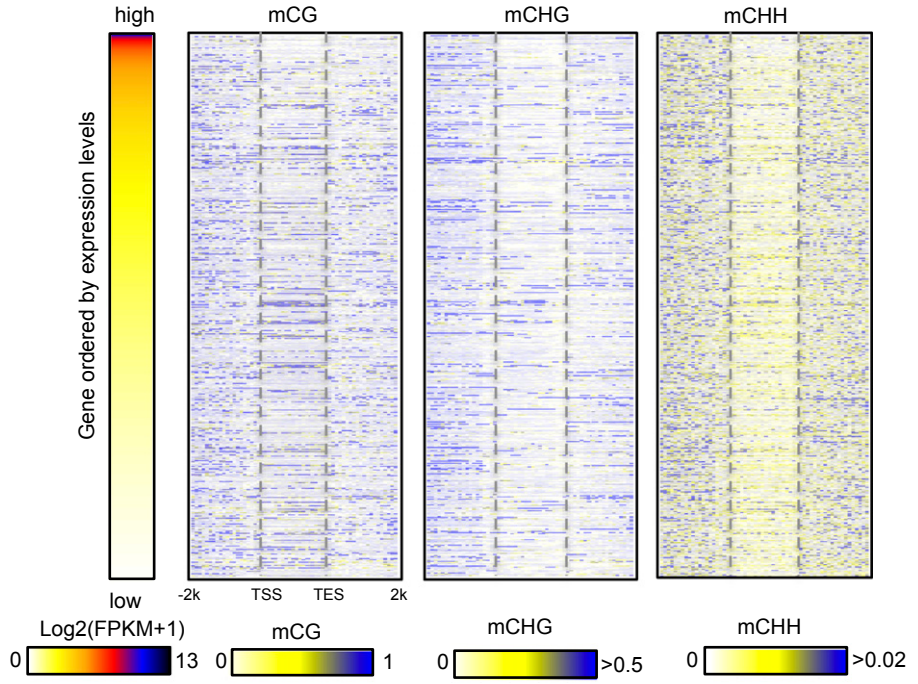

Fig. S6. Heatmap of DNA methylation in genes ordered by expression. Gene abundance was estimated by FPKM. In the case of zero value, we used log2(FPKM+1) to order gene expression. DNA methylation in each sequence context is shown correspondingly.

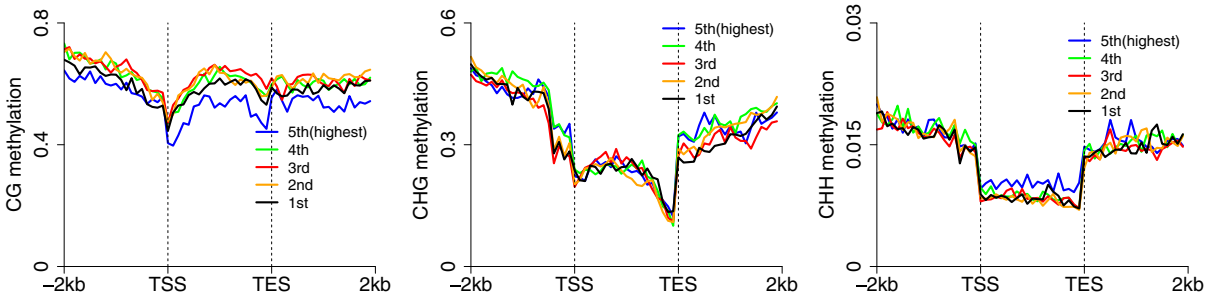

Fig. S7. Correlation between DNA methylation and transcription. Expressed genes were divided into five groups by expression level, from first (lowest expression) to fifth (highest expression). Each group contains a similar number of genes.

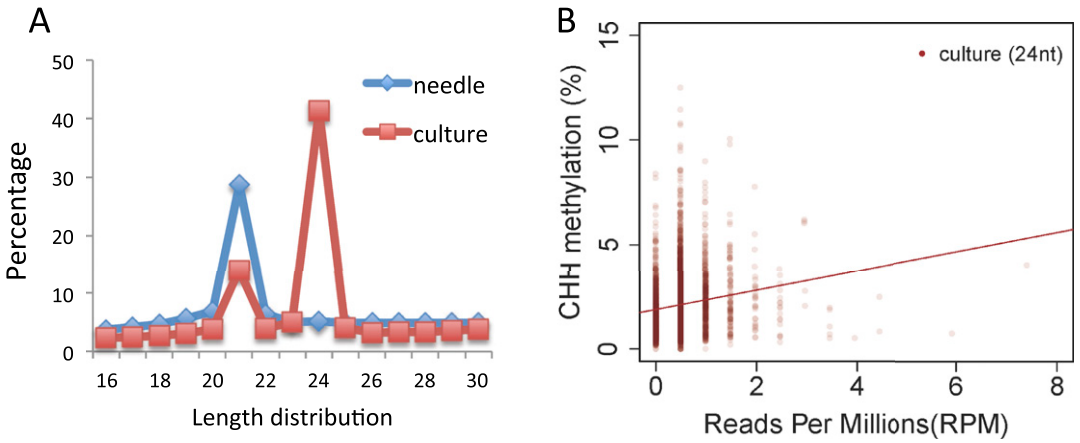

Fig. S8. Correlation between 24-nt siRNA abundance and CHH methylation levels. (*A*) Comparison of the length distributions of siRNA between needle and SE culture. (*B*) The correlation between CHH methylation and 24-nt siRNA abundance of supercontig1 in SE culture. Each bin represents a 500-bp window. Red line represents a fitted line from a linear model (lm function in R software) for SE culture.

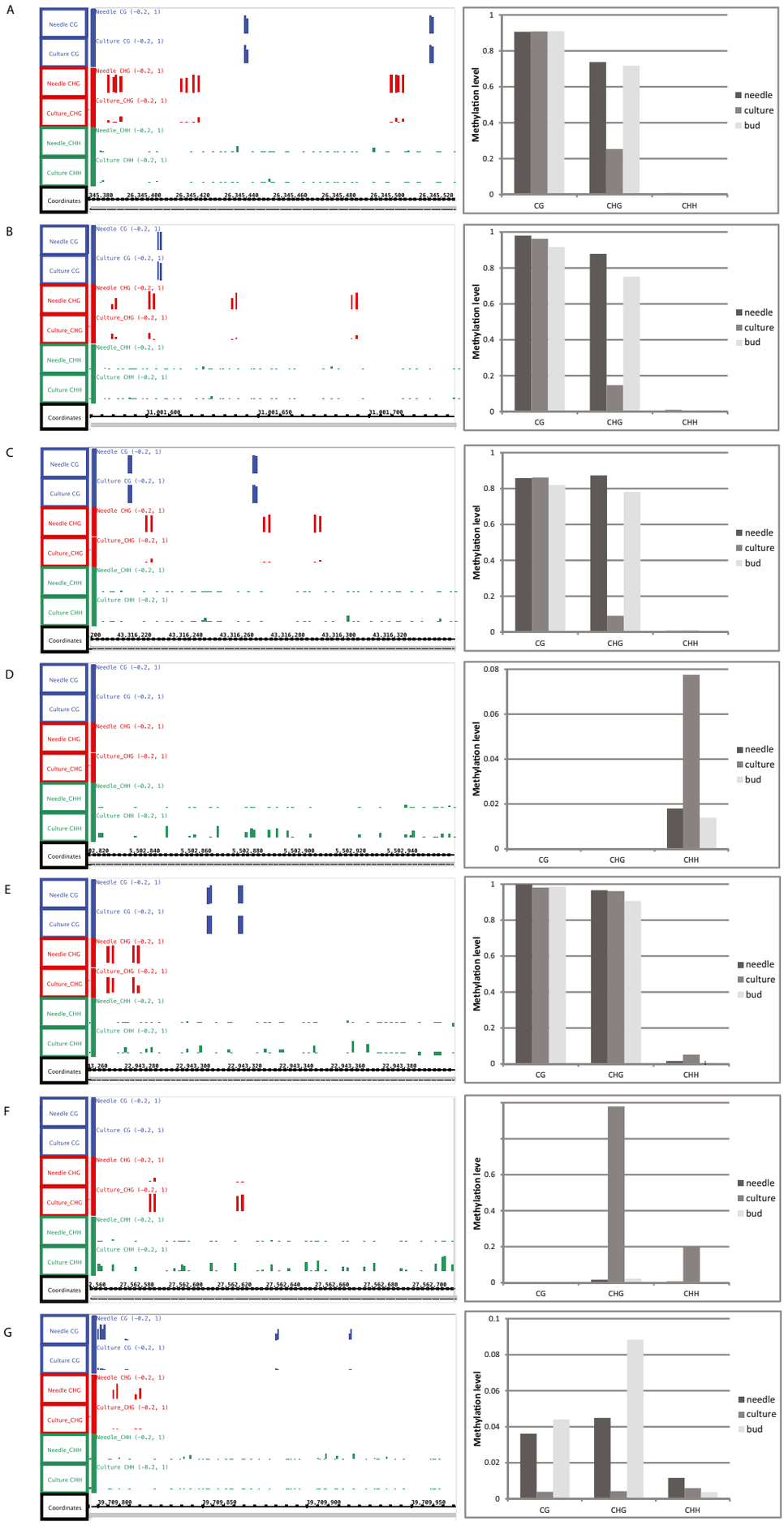

Fig. S9. Detection of methylation differences between needles and SE culture by BS-seq and validation of BS-seq data by traditional bisulfite sequencing. BS-seq data [screenshot from Integrated Genome Browser (IGB) browser] from selected regions in supercontig 1 (see *Experimental Procedures* for supercontigs) are shown on the left and traditional bisulfite data (see Table S2 for PCR primers used) are shown on the right. In BS-seq data, CG, CHG, and CHH are shown in blue, red, and green, respectively. Regions in *A*–*C* lose CHG methylation in SE culture; regions in *D* and *E* gain CHH methylation in SE culture; region in *F* gains CHG and CHH methylation in SE culture; and region in *G* loses all three types of methylation in SE culture. Traditional bisulfite sequencing data obtained from buds is also included for comparison purpose.

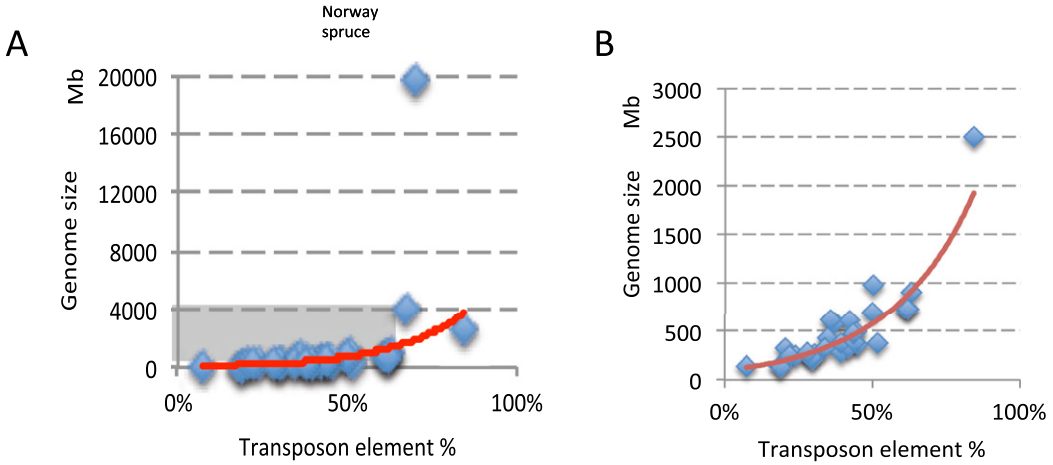

Fig. S10. Correlation between transposon element content and genome size. (*A*) Genome size is correlated with TE content. Thirty-three angiosperm species were used to show this correlation. (*B*) This is a zoomed-in view of the gray area in *A*. Red line is exponential regression line.

Table S1. Summary of BS-seq results and estimation of methylation levels

Samples	Read no.	Uniquely mapped	Mapped ratio, %	%mCG	%mCHG	%mCHH	Conversion rate, %	Sequencing depth, Gb
Needle rep1	4,071,099,559	2,374,889,504	58.3	74.7	69.1	1.5	94.91	118
Needle rep2	4,538,423,678	2,636,656,075	58.1	74.4	68.9	1.5	95.43	131
Culture rep1	3,776,335,406	2,102,719,947	55.7	68.6	62.8	2.6	94.49	105
Culture rep2	3,801,275,690	2,123,873,990	55.9	66.3	60.9	1.9	95.00	106

Table S2. PCR primers used in traditional bisulfite sequencing

Regions	Primer sequences
Supercontig1: 26,345,393–26,345,542	5′-TTGAGTGAAAAAATTYGAATATTATAAATTGTTTGGA-3 5′-TTTTAAACTTTTAACACRTTTCCATACCCT-3′
Supercontig1: 31,001,585–31,001,748	5′-YTTTTTTGAGTGTTTTTGGGTTAAATTTGAG-3′ 5′-CTAAAATAATCATTTAAAATACTTTTTCATATTCATAAATTTTATTTTTA-3′
Supercontig1: 43,316,225–43,316,367	5′-TATATTATAATTTTTTTTGTTTTATATTATGTTTTTTATTTTGYTTGAA-3′ 5′-CATCAAACAAAAATTCTCTTTRCAAAATATATAAAAAAAATAC-3′
Supercontig1: 5,502,827–5,502,968	5′-ATTTGAGTGATTGTTTTTTTTTTTTTYAGTATAYTGA-3′ 5′-CACACTARACCTATCATACCACATAATATTTC-3′
Supercontig1: 22,943,274–22,943,415	5′-TGTGYAATAATATAAYGAAATTGTGTGYGAATA-3′ 5′-TCTTATTTCTTAAAAATTTTAATTAAAACTCTCACTCCATA-3′
Supercontig1: 27,562,572–27,562,721	5′-TTATTATTGTTAATTTTTTAATTGAYGAGATTTTAATTTTTTTAATATAA-3′ 5′-AAATTCTTCCTTCAAATACAAATARARTTAAAAATTTCTTA-3′
Supercontig1: 39,709,799–39,709,973	5′-AGAGAGGATGAAGGGAATGATTGA-3′ 5′-CAATAAAAATAAAAATATAAATATTAACTAAAAATCAATACCAAAAATAAT-3′

Table S3. Summary of sRNA-seq read alignment

Samples	Total reads*	Uniquely mapped	Mapped ratio, %
Needle	2,061,322	1,494,660	72.51
SE culture	2,261,492	2,027,354	89.65

*Nonredundant reads.

Other Supporting Information Files

Dataset S1 (PDF)