Supplemental Figure 1: Efficient bisulfite conversion of non-methylated cytosine residues

(**A**) Above: Read coverage across cytosine residues from the Renilla Luciferase (R-Luc) mRNA *in vitro* transcribed BS conversion control. Below: Bisulfite conversion rates of cytosine residues in the R-Luc mRNA *in vitro* transcribed BS conversion control. (**B**) Global endogenous cytosine abundance is reduced after bisulfite treatment. Results show nucleotide (A, T, C and G) contributions from the RBS-seq libraries of the two biological replicates of wild type seedling shoots RBS-seq libraries combined. The red triangle indicates the change in cytosine abundance to less than 0.5% after bisulfite conversion transcriptome-wide. (**C**) Endogenous tRNA$^{Asp(GTC)}$ was used as a positive control to confirm previously identified m^{5}C sites at cytosine positions C38, C48, C49 and C50 and to also confirm efficient bisulfite conversion of highly structured RNAs. Results shown are from 12 individual Sanger sequenced PCR clones. Each row represents an individual clone. Methylated cytosines are shaded black, while non-methylated cytosines are white.
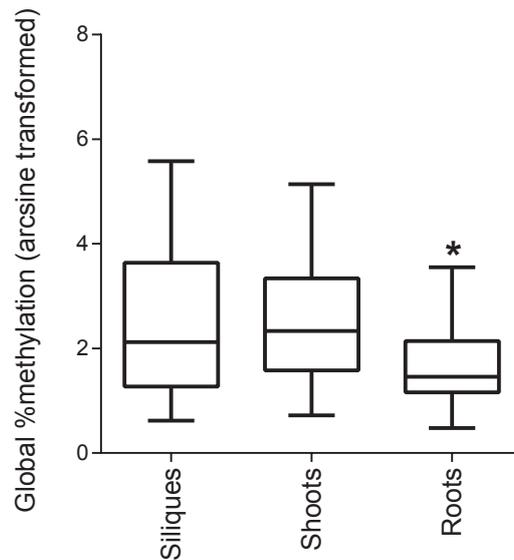
Supplemental Figure 2: Differences in global methylation levels of m⁵C sites in *Arabidopsis* siliques, shoots and roots.

Box-and-whisker plots (Tukey method) show the distributions of log2 arcsine transformed methylation percentages of transcriptome-wide m⁵C sites called in *Arabidopsis* siliques, shoots and roots (FDR≤0.3, methylation ≥1%). The bar in the box-and-whisker plots represents the median value, while the box encompasses the range of data between the first and third quartile (interquartile range). The whiskers, or error bars, show the range of data points within 1.5 times the interquartile range. Differences between global methylation levels in the three tissue types were subjected to one way ANOVA and Tukey's multiple comparisons test (*P-value ≤0.0001).

A

B

TRM4A

| Motif X | Motif IV | Motif VI | Motif VIII |

TRM4B

| Motif X | Motif I | Motif IV | Motif VI | Motif VIII |
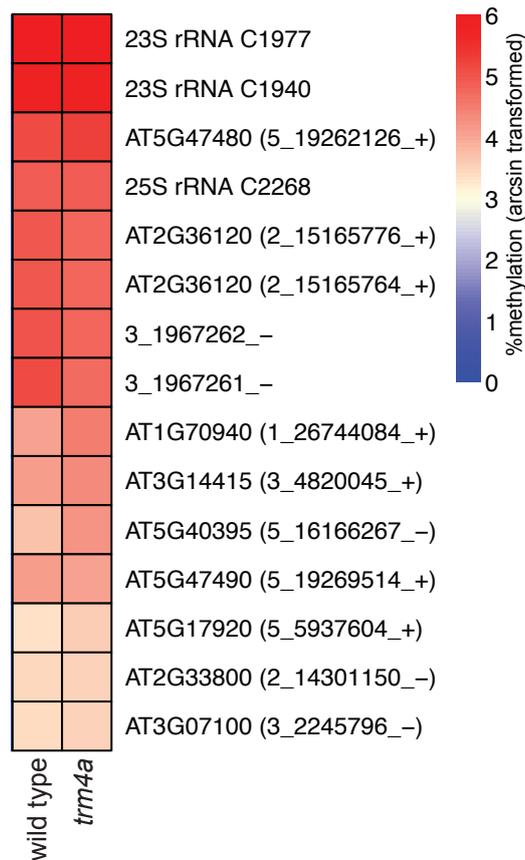
Supplemental Figure 3: Methylation of selected m⁵C sites analyzed using bsRNA-amp-seq are not perturbed in *trm4a* mutants.

(**A**) Heat map showing percentage methylation of cytosines derived from targeted bsRNA-amp-sequencing in wild type and *trm4a* seedling shoots. Cytosine positions are indicated on the left (3 biological replicates). (**B**) The amino acid sequences of *Arabidopsis* RNA 5-methylcytosine methyltransferases TRM4A and TRM4B were compared and the methyl-transferase motifs were identified and are shown as labeled purple boxes in the schematic. TRM4A is missing Motif I, which has previously been shown to be critical for RMTase activity.
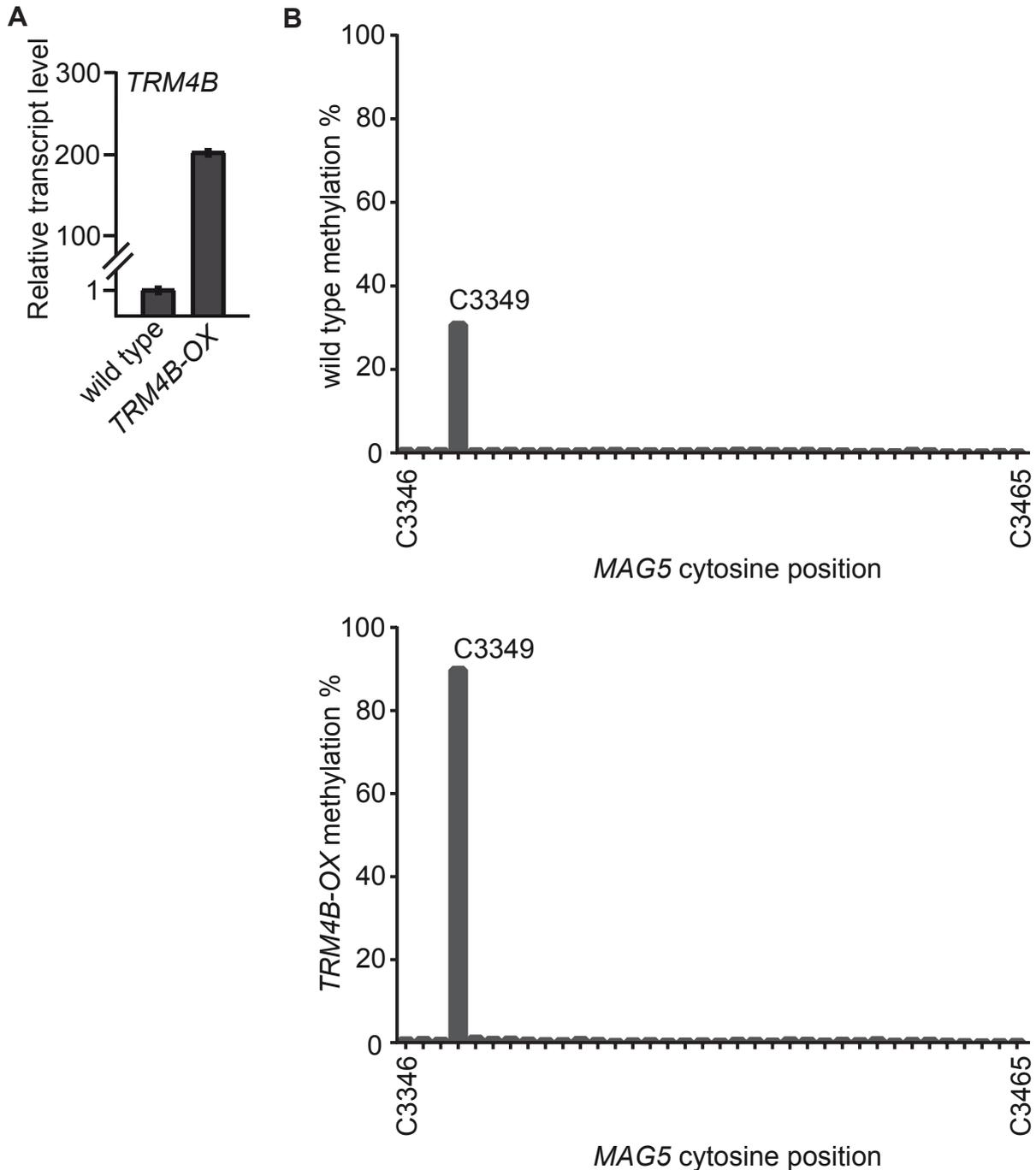
Supplemental Figure 4: qRT-PCR expression of *TRM4B-OX* line and BS conversion efficiency of bsRNA-amplicon-seq.
(**A**) Quantitative RT-PCR of *TRM4B* transcripts in wild type and *TRM4B-OX*. Expression was normalized to *PDF2A* mRNA abundance. Error bars represent s.e. (n=3 technical replicates). (**B**) Methylation percentages across the 36 cytosine residues from the *MAG5* region amplified using bsRNA-amp-seq in wild type (above) and *TRM4B-OX* (below). The TRM4B dependent m^{5}C site at position C3349 shows increased methylation in plants over-expressing *TRM4B* compared to wildtype and no increases in methylation at neighboring cytosines.
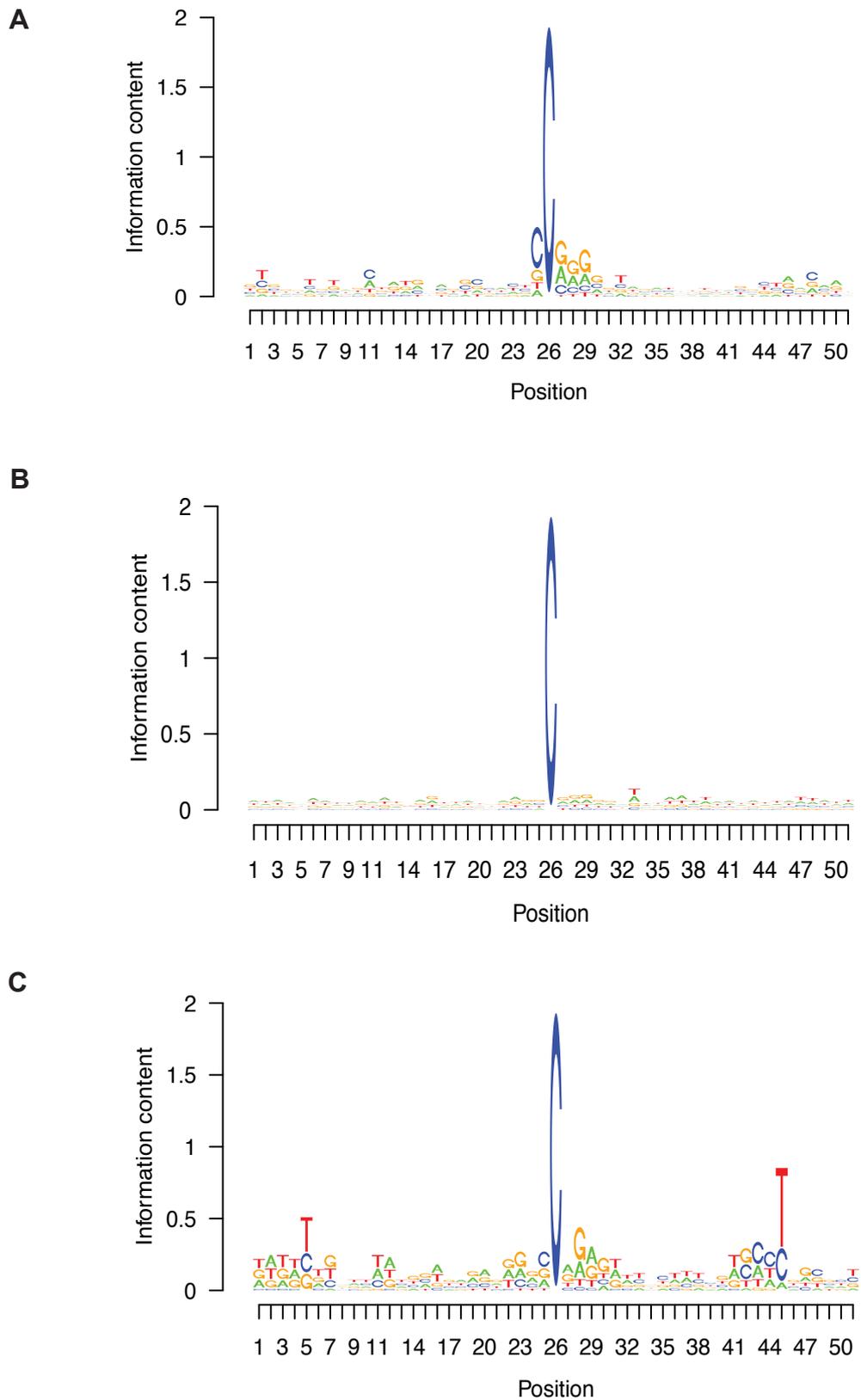
Supplemental Figure 5: LOGO motif analysis of TRM4B dependent m⁵C sites.
LOGO motif analysis of TRM4B dependent sites from *Arabidopsis* (**A**) siliques (**B**) shoots and (**C**) roots.
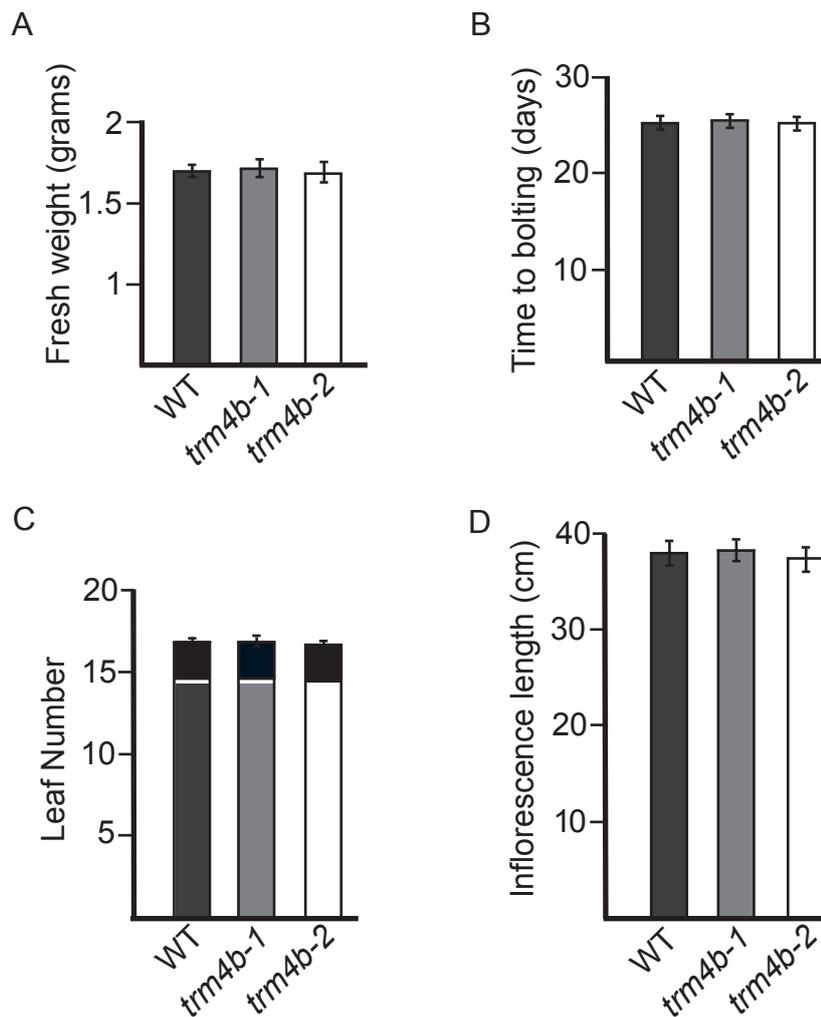
Supplemental Figure 6: Vegetative and flowering time traits of wild type and *trm4b* mutants.
(A) Fresh weight of wild type (WT) and *trm4b* shoots at primary inflorecsence emergence
(n=10). (B) Flowering time of WT and *trm4b* plants was measured and is shown as days to appearance of the primary inflorescence (n=16). (C) Flowering time of the plants in panel b shown as rosette leaf (lower section of bar) and cauline leaf number (upper black section, n=16). (D) Primary inflorescence length was measured of the plants in panel b 16 days after primary infloresence emergence. In all panels, error bars represent SE of the mean. No statistical difference (P<0.05, Students t-test) was observed between WT and mutant plants for the four traits measured.
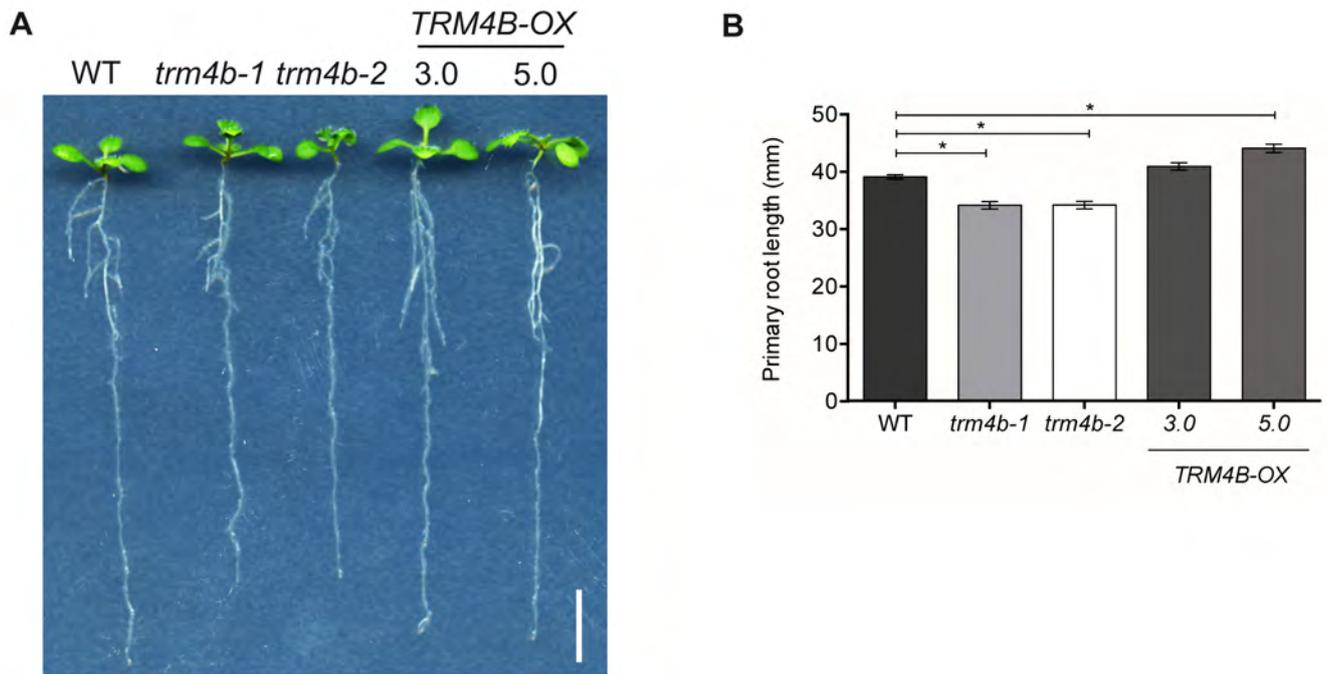
Supplemental Figure 7: Complementation of *trm4b* short-root phenotype.

(A) Root elongation of seven day after germination seedlings of wild type (Col-0); *trm4b-1* and *trm4b-2* T-DNA mutants and two independent TRM4B over-expression transgenic lines; 3.0 and 5.0, grown on half-strength MS medium with 1% sucrose. Scale bar, 1 cm. (B) Quantification of primary root length (7 DAG) of wild type, trm4b mutants and TRM4B-OX lines. Error bars represent SE of the mean (*P<0.01, Students t-test; n= 10 seedlings).

A

B
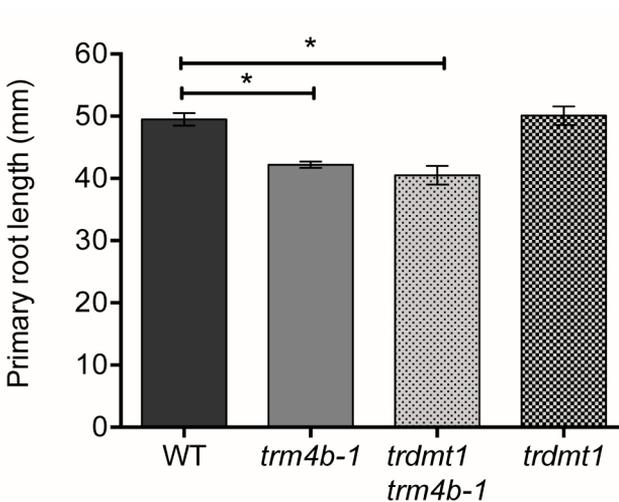
Supplemental Figure 8: Primary root length of *trdmt1* mutants is similar to wild type plants.
(A) Root elongation of 10 days after germination (DAG) wild type, trdmt1, *trm4b-1* and *trdmt1 trm4b* double mutants, grown on halfstrength MS media with 1% sucrose. Scale bar, 1 cm. (B) Primary root growth measurements at 10 DAG for wild type, *trm4b-1, trdmt1 trm4b* and *trdmt1* mutants. Error bars indicate SE of the mean (*P<0.05, Students t-test; n≥14 seedlings)

A

B
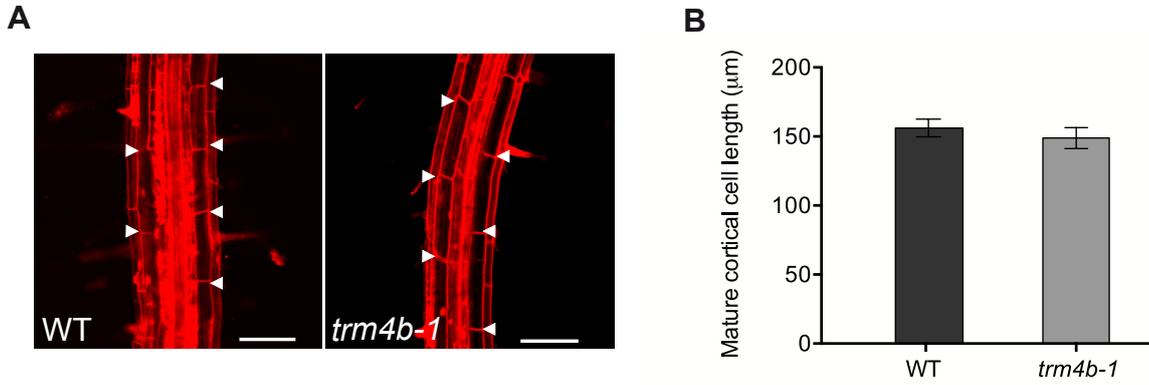
Supplemental Figure 9: Mature cortical cell length of wild type and *trm4b-1* mutant roots.
(A) Propidium iodide stained confocal images of cortical cells in the root differentiatiation zone of wild type and *trm4b-1* mutant seedlings (7 DAG). White arrowheads indicate cortical cell wall boundaries. Scale bar, 100 µm. (**B**) Average mature root cortical cell length for wild type and *trm4b-1* mutant. Error bars indicate the SE of the mean. There was no difference in mature cortical cell length as determined by Students t-test (P<0.05); n = 9 seedlings).

A

WT   *trm4b-1*

His^{GTG}

1.0   1.0   0.8   0.7

EtBr rRNA

1.0   1.0   1.0   1.0

B
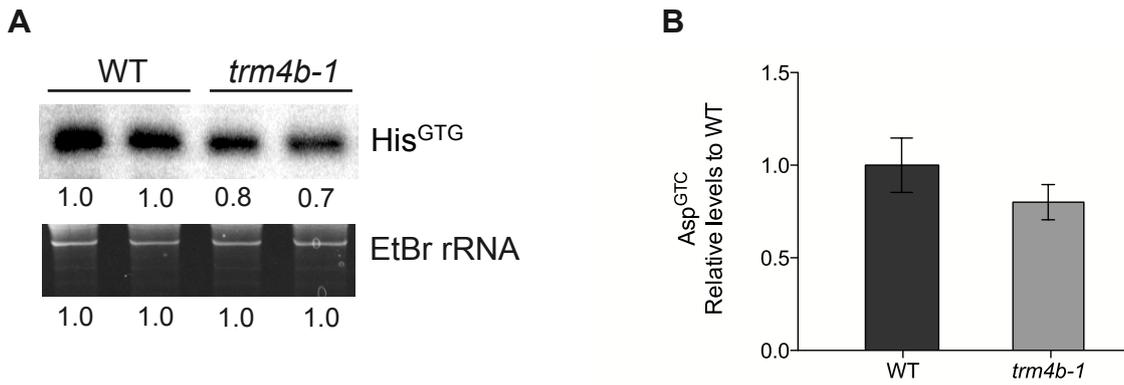
Supplemental Figure 10: RNA gel blot analysis of tRNA$^{His(GTG)}$ **.** (**A**) RNA gel blot detection of TRM4B substrate tRNA$^{His(GTG)}$ in wild type and *trm4b-1*. Normalised intensity values are given beneath each lane. (B) Average signal intensity of tRNA$^{His(GTG)}$ in wild type and *trm4b-1*. The tRNA signal intensities were normalized to rRNA loading control and then compared to wild type. Error bars represent SE of the mean; n=2 replicate RNA samples from one RNA gel blot experiment.

A

single–organism process
single–organism cellular process
★ response to stimulus
★ response to chemical
localization
single–organism metabolic process
single–organism biosynthetic process
cellular protein localization
response to nitrogen compound
small molecule metabolic process
cellular process
response to endogenous stimulus
immune system process
response to abiotic stimulus
primary metabolic process
innate immune response
★ response to oxidative stress
organic hydroxy compound biosynthetic process
★ cell differentiation
phosphorus metabolic process
response to external stimulus
organic hydroxy compound metabolic process
organic cyclic compound biosynthetic process
metabolic process
organic cyclic compound metabolic process
ncRNA processing
nitrogen compound transport
benzene–containing compound metabolic process
organophosphate biosynthetic process
response to ethylene
★ hydrogen peroxide catabolic process
biological regulation
ncRNA metabolic process
aromatic compound biosynthetic process
cellular aromatic compound metabolic process
response to biotic stimulus
regulation of biological process
organonitrogen compound metabolic process
monocarboxylic acid metabolic process
carbohydrate derivative metabolic process
response to external biotic stimulus
★ reactive oxygen species metabolic process
cell communication
regulation of response to stimulus
ion homeostasis
nitrogen compound metabolic process
★ oxidation–reduction process
multi–organism process
macromolecule modification
single–organism membrane organization
membrane organization
cellular metabolic process
response to carbohydrate
organic substance metabolic process
secondary metabolic process
RNA processing
response to light stimulus
lipid biosynthetic process
cellular component organization or biogenesis
rhythmic process
developmental process
lipid metabolic process
catabolic process
secondary metabolite biosynthetic process
carbohydrate derivative biosynthetic process
pyrimidine–containing compound metabolic process
pigment metabolic process
circadian rhythm
protein metabolic process
heterocycle metabolic process
cellular nitrogen compound metabolic process
oligopeptide transport
response to wounding
carbohydrate metabolic process
chlorophyll metabolic process
negative regulation of programmed cell death
response to fructose
developmental maturation
negative regulation of biological process
pigment biosynthetic process
photosynthesis, light reaction
signal transduction by phosphorylation
multicellular organismal process
★ cellular response to hypoxia
★ respiratory burst
positive regulation of metabolic process
MAPK cascade
cellular carbohydrate metabolic process
organic acid transport
cellular biogenic amine catabolic process
regulation of multicellular organismal process
flavonoid biosynthetic process
generation of precursor metabolites and energy
regulation of phosphate metabolic process
regulation of phosphorus metabolic process
dephosphorylation
positive regulation of biological process
plastid organization
positive regulation of catalytic activity
flavonoid metabolic process
RNA modification
regulation of protein modification process
response to transition metal nanoparticle

−log$_{10}$(p-value)

50

40

30

20

10

paraquat WT
vs control WT

B

root epidermal cell differentiation
single–organism process
single–organism cellular process
★ response to stimulus
response to nitrogen compound
phenylpropanoid metabolic process
developmental process
response to nitrate
secondary metabolic process
localization
cell wall organization or biogenesis
anion transport
★ response to chemical
developmental maturation
single–organism metabolic process
cell wall organization
cellular response to iron ion starvation
protein localization to nucleus
nitrogen compound transport
pyrimidine–containing compound metabolic process
pyrimidine ribonucleotide biosynthetic process
★ respiratory burst
★ respiratory burst involved in defense response
cullin deneddylation
polysaccharide metabolic process
cellular localization
cellular process
response to red or far red light
★ oxidation–reduction process
cell communication
response to extracellular stimulus
cell junction organization
cell–cell junction organization
single–organism biosynthetic process
★ response to oxidative stress

−log$_{10}$(p-value)

16

14

12

10

8

6

4

control *trm4b-1*
vs control WT

Supplemental Figure 11: Oxidative stress responsive biological process GO terms are constitutively activated in *trm4b* mutants. Whole roots at 6 DAG were harvested from wild type (WT), *trm4b-1* and paraquat treated wild type seedlings and RNA-seq analysis was performed (n=3). Biological process GO terms involved in the oxidative stress response are significantly enriched (FDR<0.05) in (**A**) paraquat treated wild type plants compared to wild type controls and in (**B**) *trm4b-1* mutants compared to wild type plants. Black stars indicate GO terms relating to oxidative stress responses. The heat map shows the significance level using the negative log of the p-value, where red = very significant and yellow = significant (FDR <0.05).

Supplemental Data Set Legends:

Supplemental Data Set 1: Read coverage of all libraries sequenced
Table summarizes the total number of reads obtained using bsRNA-seq or RNA-seq of *Arabidopsis* wild type and mutant genotypes used in this study. The tissue and/or treatment used and the number of biological and technical replicates for each library sample is also indicated.

Supplemental Data Set 2: Overall m⁵C site calls

Supplemental Data Set 2: Overall m⁵C site calls
The table details individual m⁵C sites identified from bsRNA-seq of *Arabidopsis* silique, shoot and root transcriptomes and is organised in three separate worksheets respectively. The genomic location of individual sites, read number, methylation percentage and associated p-value, adjusted p-value and FDR is specified for each replicate library sequenced. Average methylation percentage values that meet the statistical cut-offs (FDR ≤0.3 and ≥1% methylation) as described in the Materials and Methods section are specified for each m⁵C site. For all m⁵C sites that are located within protein coding genes, the gene symbol and description along with gene feature category are listed.

Supplemental Data Set 3: *trm4b* dependent m⁵C sites
The table details *trm4b* dependent m⁵C sites identified from bsRNA-seq of *Arabidopsis* silique, shoot and root transcriptomes and is organised in three separate worksheets respectively. The genomic location of individual sites, read number, methylation percentage and associated p-value, adjusted p-value and FDR is specified for each replicate library sequenced. Average methylation percentage values that meet the statistical cut-offs (FDR ≤0.3 and ≥1% methylation) as described in the Materials and Methods section are specified for each m⁵C site. For all m⁵C sites that are located within protein coding genes, the gene symbol and description along with gene feature category are listed.

Supplemental Data Set 4: wild type heat map raw values Fig.1g
The table provides the raw methylation values of five m⁵C sites determined using transcriptome-wide bsRNA-seq and targeted bsRNA-amp-seq approach represented as a heat map in Figure 1G. The average methylation percentage for each site (before and after arc sin transformation) is listed for silique, shoot and root tissues analysed in this study.

Supplemental Data Set 5: *trdmt1* dependent m5C sites
The table provides m⁵C methylation levels of all sites identified in *Arabidopsis trdmt1* mutant and wild type seedling shoot transcriptomes. No m⁵C sites were identified that were significantly reduced in *trdmt1* mutant compared to wild type (FDR ≤0.3).

Supplemental Data Set 6: *trm4a* heat map raw values Supplementary Fig. 3a

Raw methylation values of fifteen m⁵C sites in *Arabidopsis trm4a* mutant and wild type represented as heat map in supplementary Figure 3a. The average methylation percentage for each site (before and after arc sin transformation) as determined using bsRNA-amp-seq is specified.

Supplemental Data Set 7: *trm4b* heat map raw values Fig. 3D
Raw methylation values of *trm4b* dependent candidate m⁵C sites independently validated using bsRNA-amp-seq approach represented as a heat map in Figure 3D. The average methylation percentage for each site (before and after arc sin transformation) is listed for silique, shoot and root tissues analysed in this study.

Supplemental Data Set 8: TRM4B-OX raw values for heat map Fig. 3E
Raw methylation values of *trm4b* dependent m⁵C sites in wild type and transgenic plants overexpressing *TRM4B* (*TRM4B-OX*) represented as a heat map in Figure 3E. The methylation percentage of the ten sites in wild-type and *TRM4B-OX* was determined using bsRNA-amp-seq approach.

Supplemental Data Set 9: RNA-seq analysis of Differentially Expressed genes in wild type and trm4b mutant roots under controlled and paraquat stressed conditions
This table summarises the RNA-seq differential expression analysis of the *trm4b* mutant and wild type in paraquat stressed and control conditions from *Arabidopsis* whole root tissue. The differentially expressed genes for each comparison are organised in separate worksheets and include *trmb4b* mutant versus wild type in control conditions (R_ctrl_trm4b__ctrl_WT), *trmb4b* mutant versus wild type in paraquat stressed conditions (R_para_trm4b__para_WT), paraquat stressed wild type versus control wild type (R_para_WT__ctrl_WT) and paraquat stressed *trmb4b* mutant versus control *trm4b* mutant (R_para_trm4b__ctrl_trm4b). Genes that showed a significant interaction between the loss of TRM4B and oxidative stress from the RNA-seq data (FDR ≤0.05) are listed on worksheet "R_interaction_terms". For each differential expressed gene, we report the gene ID; chromosomal location; transcript length; log Fold Change (logFC) and statistical cut off to measure significance (p-value, adjusted p-value and FDR).

Supplemental Data Set 10: Primer sequences used in the study.