New Phytologist

***New Phytologist* Supporting Information**

Article title: Deciphering the synergistic and redundant roles of CG and non-CG DNA methylation in plant development and TE silencing
Authors: Wenjie Liang, Jinchao Li, Linhua Sun, Yi Liu, Zijun Lan, and Weiqiang Qian
Article acceptance date: 11 October 2021

The following Supporting Information is available for this article:

Fig. S1 Generation of *MET1* mutants in *Arabidopsis* using the CRISPR/Cas9 system.

Fig. S2 Loss of DNA methylation affects endosperm development and compromises transmission through the male gametophytes in *Arabidopsis*.

Fig. S3 Loss of DNA methylation does not affect pollen viability in *Arabidopsis*.

Fig. S4 Loss of DNA methylation in *Arabidopsis* leads to the downregulation of some genes involved in pollen development and double fertilization in mature pollens.

Fig. S5 Loss of DNA methylation leads to global changes of gene expression in *Arabidopsis* seedlings.

Fig. S6 The upregulation of cell cycle-, DNA replication- and DNA repair-related genes in *Arabidopsis ddcc met1-11*$^{+/-}$ seedlings is not due to loss of DNA methylation in the promoters of these genes.

Fig. S7 The *Arabidopsis ddcc met1-11*$^{+/-}$ mutation specifically activates a group of TEs.

Fig. S8 CG and non-CG DNA methylation synergistically regulate chromocenter condensation in *Arabidopsis*.

Table S1 Primers and sgRNAs used in this study.

Tables S2–S5 (see separate files)

Fig. S1 Generation of *met1* mutants in *Arabidopsis* using the CRISPR/Cas9 system. (a) Schematic representation of *MET1* gene structure and the mutation sites. The sgRNA targeting sequences are highlighted in yellow and the PAM sites are underlined. The *met1-10* mutation causes a five nucleotide deletion in the first exon (highlighted in red). The *met1-11* mutation causes a 'T' insertion in the sixth exon (highlighted in red). Each of these mutations causes a shift in open reading frame. **(b)** Schematic diagram of mutant screening. The CRISPR constructs were transformed into the wild-type Col-0 or *ddcc* mutants (T0). The *met1* chimera plants in T1 generation were identified through DNA sequencing, and then backcrossed to Col-0 and *ddcc*, respectively, to obtain $met1^{+/-}$ $Cas9^{-/-}$ and *ddcc* $met1^{+/-}$ $Cas9^{-/-}$ plants in F1 generation. The $met1^{-/-}$ $Cas9^{-/-}$ and *ddcc* $met1^{+/-}$ $Cas9^{-/-}$ plants in F2 generation were finally identified for further experiments. The most important plant materials in each generation are highlighted by red circles.

New Phytologist

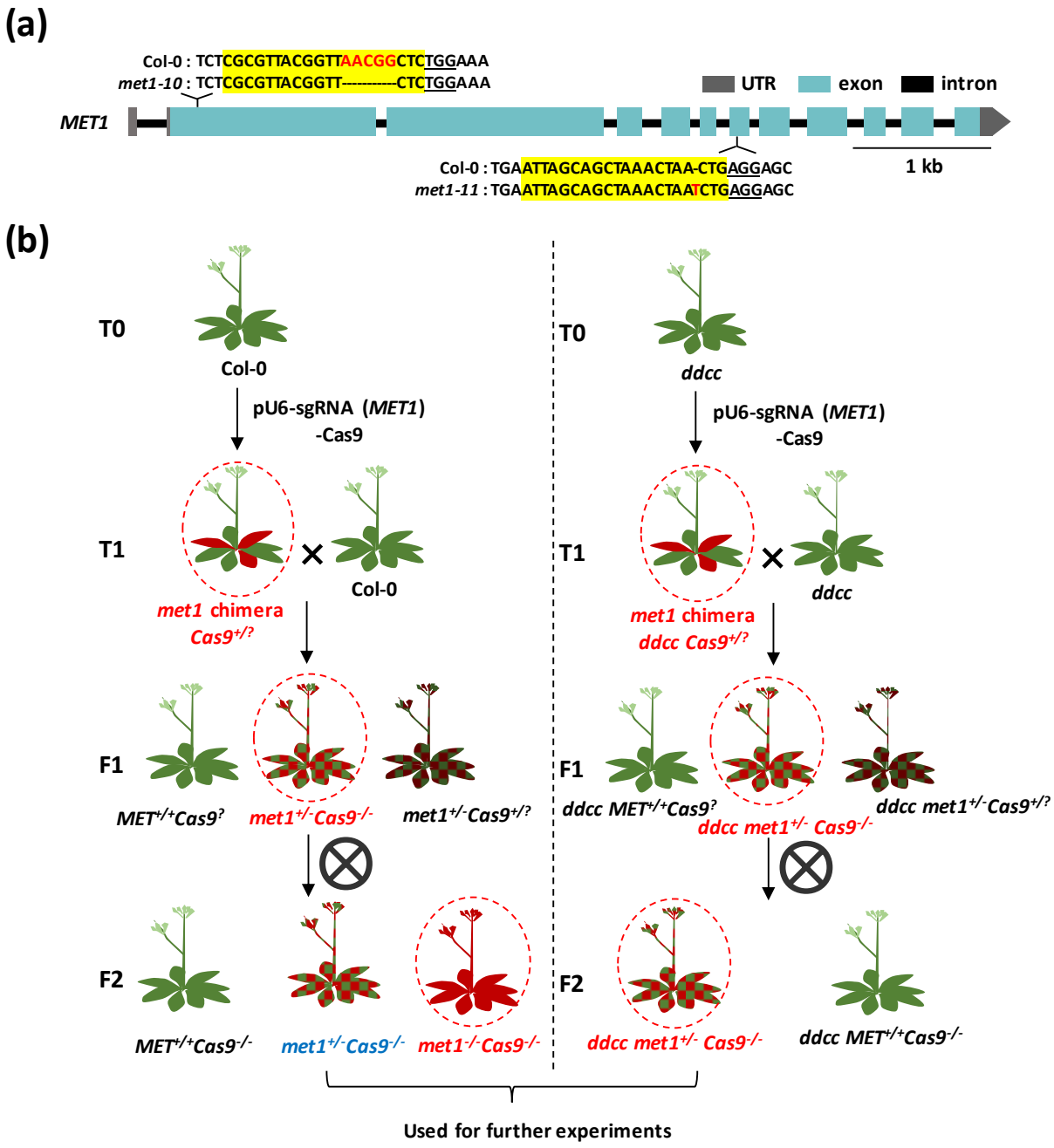
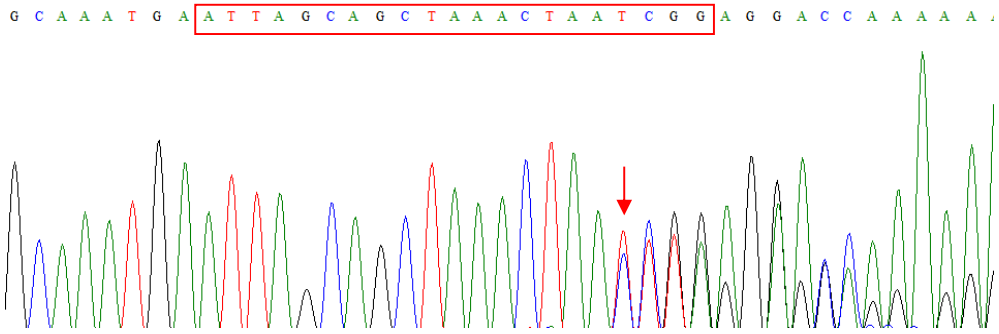
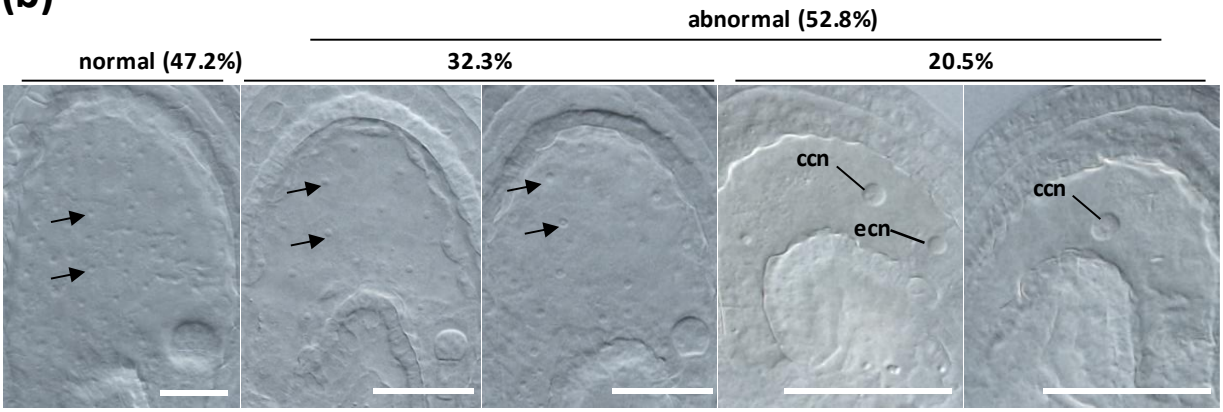


Fig. S2 Loss of DNA methylation affects endosperm development and compromises transmission through the male gametophytes in *Arabidopsis*. **(a)** DNA sequencing results showing genotypes of late aborted embryos in *ddcc met1-11*$^{+/-}$ siliques. Red box represents the location of sgRNA, and red arrow indicates the mutation site. **(b)** DIC microscopy of endosperm in *ddcc met1-11*$^{+/-}$ siliques at 3 d after pollination (DAP; n=127). The percentages of seeds with each type of endosperm are indicated. Black arrow indicates endosperm nucleus. ccn, central cell nucleus; ecn, egg cell nucleus. Bars, 50 μm. **(c)** Seed set of hand-pollinated plants (6 DAP). Red and blue arrows indicate early- and late-aborted seeds, respectively. Bars, 1 mm.

(a)


G C A A A T G A A T T A G C A G C T A A A C T A A T C G G A G G A C C A A A A A

(b)


abnormal (52.8%)
normal (47.2%)
32.3%
20.5%
ccn
ecn
ccn

(c)

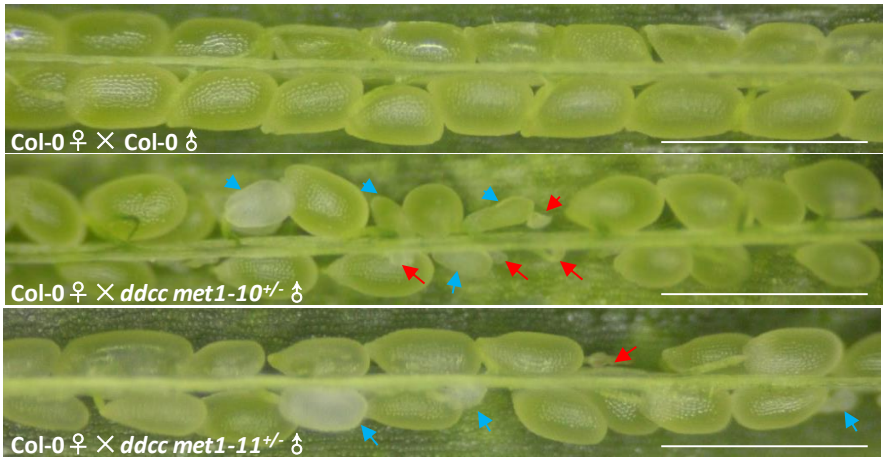
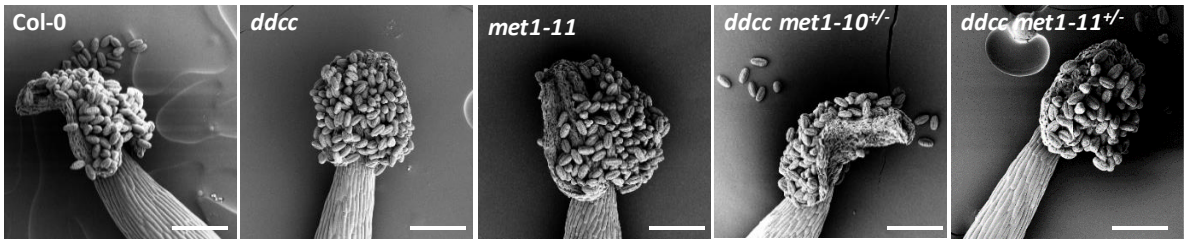
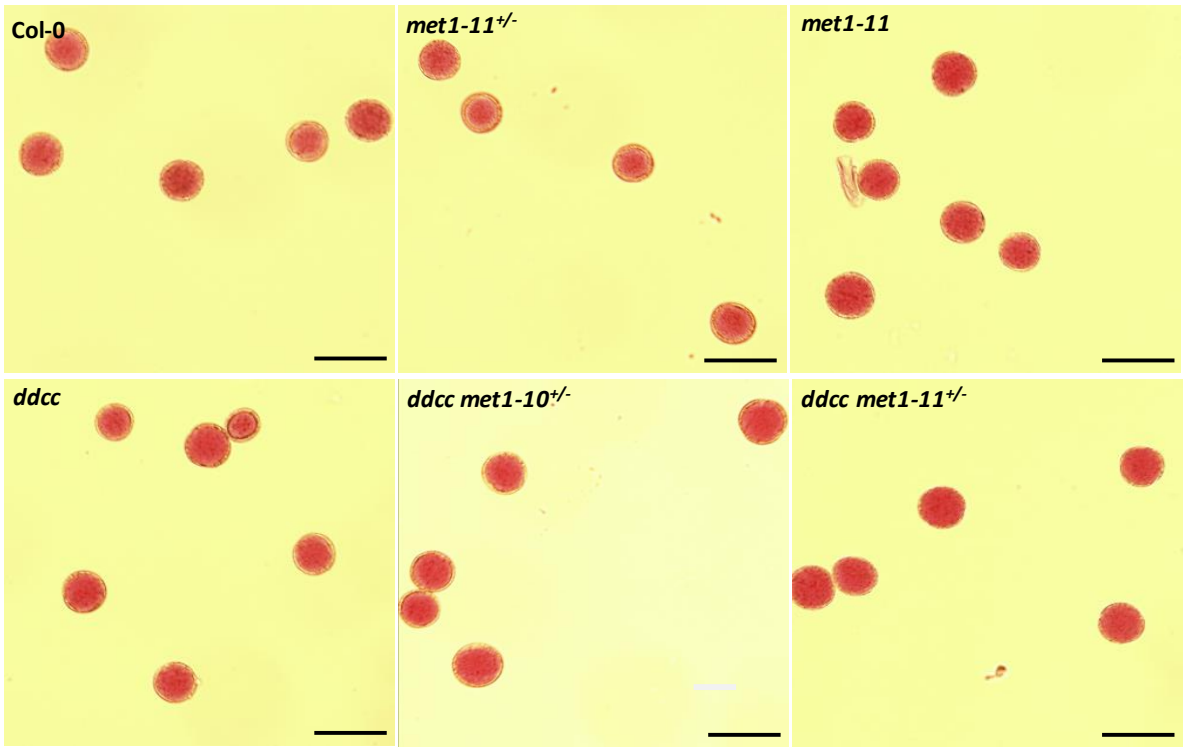

Col-0 ♀ × Col-0 ♂
Col-0 ♀ × ddcc met1-10+/- ♂
Col-0 ♀ × ddcc met1-11+/- ♂

Fig. S3 Loss of DNA methylation does not affect pollen viability in *Arabidopsis*. (a) Anther morphology of Col-0, *met1-11*, *ddcc*, *ddcc met1-10*$^{+/-}$, and *ddcc met1-11*$^{+/-}$ plants as observed using cryo-SEM. Bars, 100 μm. **(b)** Pollen activity of Col-0, *met1-11*$^{+/-}$, *met1-11*, *ddcc*, *ddcc met1-10*$^{+/-}$, and *ddcc met1-11*$^{+/-}$ plants as determined by Alexander dye staining. Bars, 100 μm. **(c)** Percentages of normal and phenotypically abnormal pollens observed in Col-0, *met1-11*$^{+/-}$, *met1-11*, *ddcc,* and *ddcc met1*$^{+/-}$ mutant plants. For each genotype, at least 1000 pollen grains were examined.

(a)


Col-0
ddcc
met1-11
ddcc met1-10+/-
ddcc met1-11+/-

(b)


Col-0
met1-11+/-
met1-11
ddcc
ddcc met1-10+/-
ddcc met1-11+/-

(c)

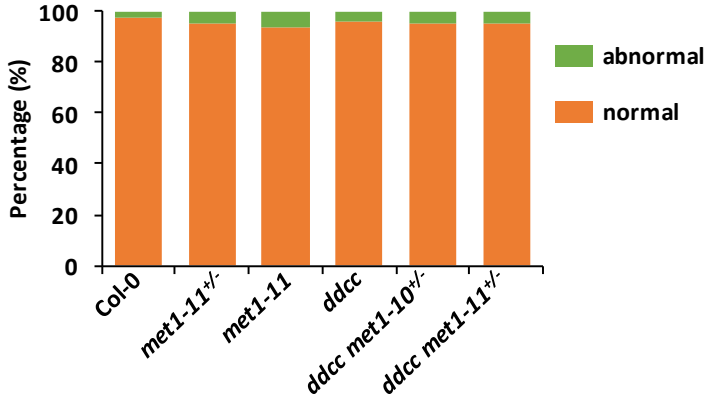
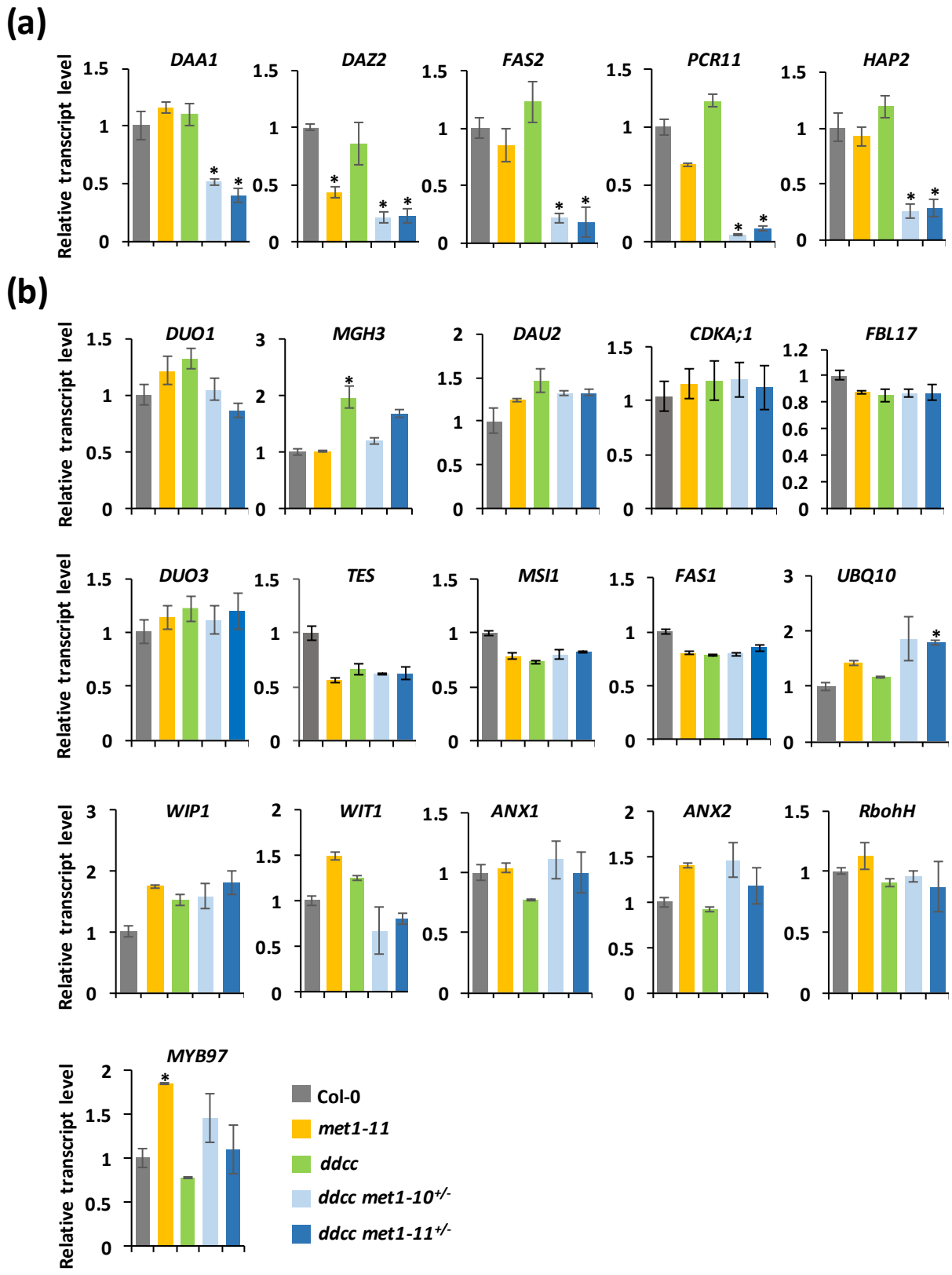

100
80
60
40
20
0
Percentage (%)
abnormal
normal
Col-0
met1-11+/-
met1-11
ddcc
ddcc met1-10+/-
ddcc met1-11+/-

Fig. S4 Loss of DNA methylation in *Arabidopsis* leads to the downregulation of some genes involved in pollen development and fertilization in mature pollens. (a) Transcript levels of genes that are significantly down-regulated in mature pollen of the *ddcc met1*$^{+/-}$ plants. Data represent the mean ± SD of three biological replicates. Asterisks indicate significant differences between Col-0 and the indicated mutants ($P < 0.05$, two-tailed Student's *t*-test). **(b)** Transcript levels of genes that have little change in mature pollens of the *ddcc met1*$^{+/-}$ plants. *DUO1*, *MGH3, DAU2, CDKA;1, FBL17, DOU3, TES, MSI1,* and *FAS1* are specifically expressed in the sperm cell; *UBQ10*, *WIP1,* and *WIT1* are specifically expressed in the vegetative cell; *ANX1*, *ANX2*, *RbohH,* and *MYB97* are genes that regulate pollen tube bursting. Data represent the mean ± SD of three biological replicates. Asterisks indicate significant differences between Col-0 and the indicated mutants ($P < 0.05$, two-tailed Student's *t*-test). **(c)** Snapshot in the Integrated Genome Browser showing the levels of DNA methylation at loci *DAA1*, *DAZ2*, *FAS2, PCR11,* and *HAP2* in the indicated genotypes. *big*: *ddcc met1-11*$^{+/-}$*-big*; *small*: *ddcc met1-11*$^{+/-}$*-small.*


(a)
DAA1
DAZ2
FAS2
PCR11
HAP2
Relative transcript level
(b)
DUO1
MGH3
DAU2
CDKA;1
FBL17
DUO3
TES
MSI1
FAS1
UBQ10
WIP1
WIT1
ANX1
ANX2
RbohH
MYB97
Col-0
met1-11
ddcc
ddcc met1-10+/-
ddcc met1-11+/-

(c)

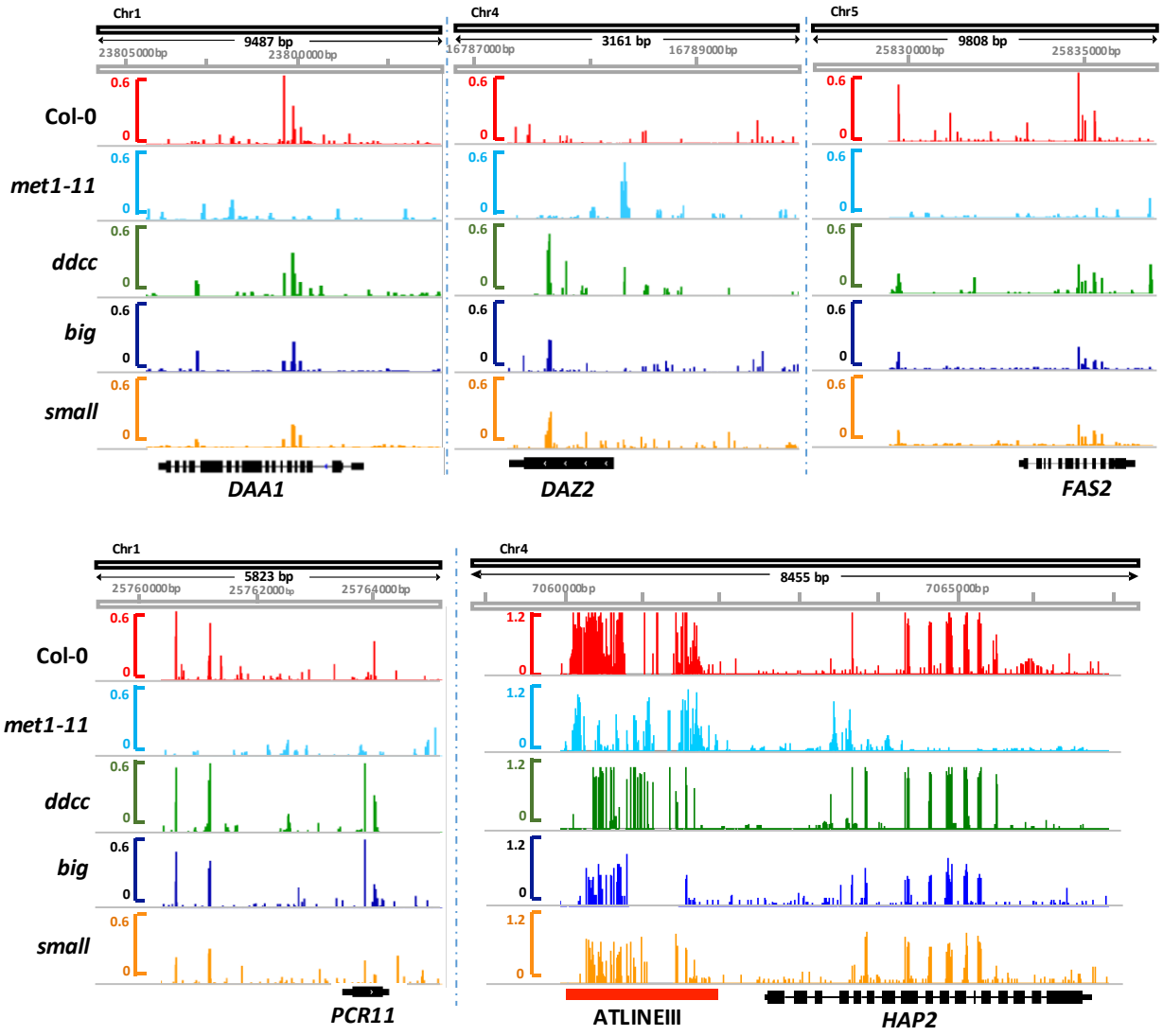
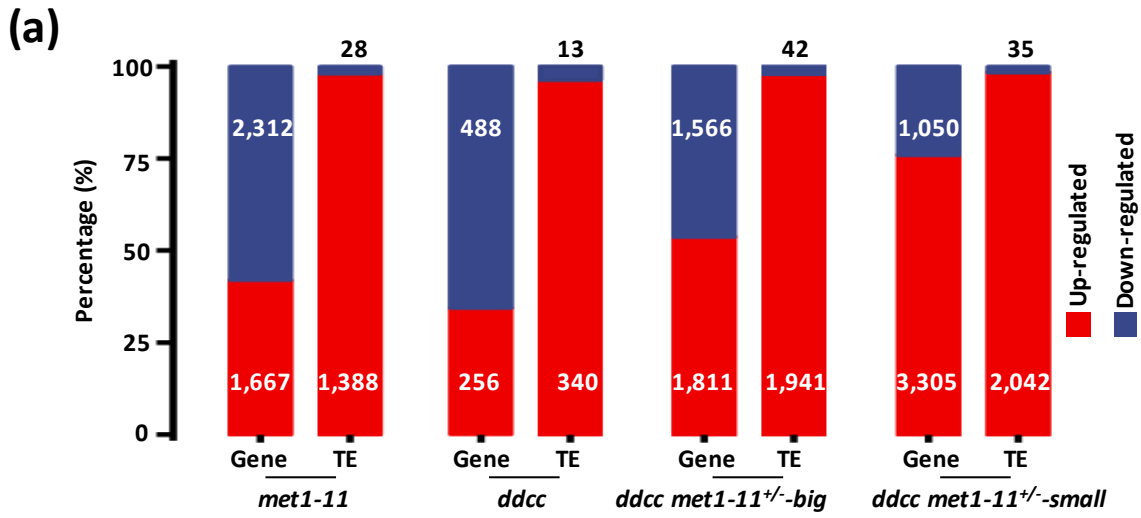

Chr1
9487 bp
23805000bp
23800000bp
Chr4
3161 bp
16787000bp
16789000bp
Chr5
9808 bp
25830000bp
25835000bp
Col-0
met1-11
ddcc
big
small
0.6
0
DAA1
DAZ2
FAS2
Chr1
5823 bp
25760000bp
25762000bp
25764000bp
Chr4
8455 bp
7060000bp
7065000bp
1.2
PCR11
ATLINEIII
HAP2

Fig. S5 Loss of DNA methylation leads to global changes of gene expression in *Arabidopsis* seedlings. (a) Up- and down-regulated genes and transposable elements (TEs) in seedlings of different genotypes. **(b)** Overlap of up- and down-regulated genes and TEs between the 'big' and 'small' plants of *ddcc met1-11*$^{+/-}$. The overlap was tested using Fisher's exact test.



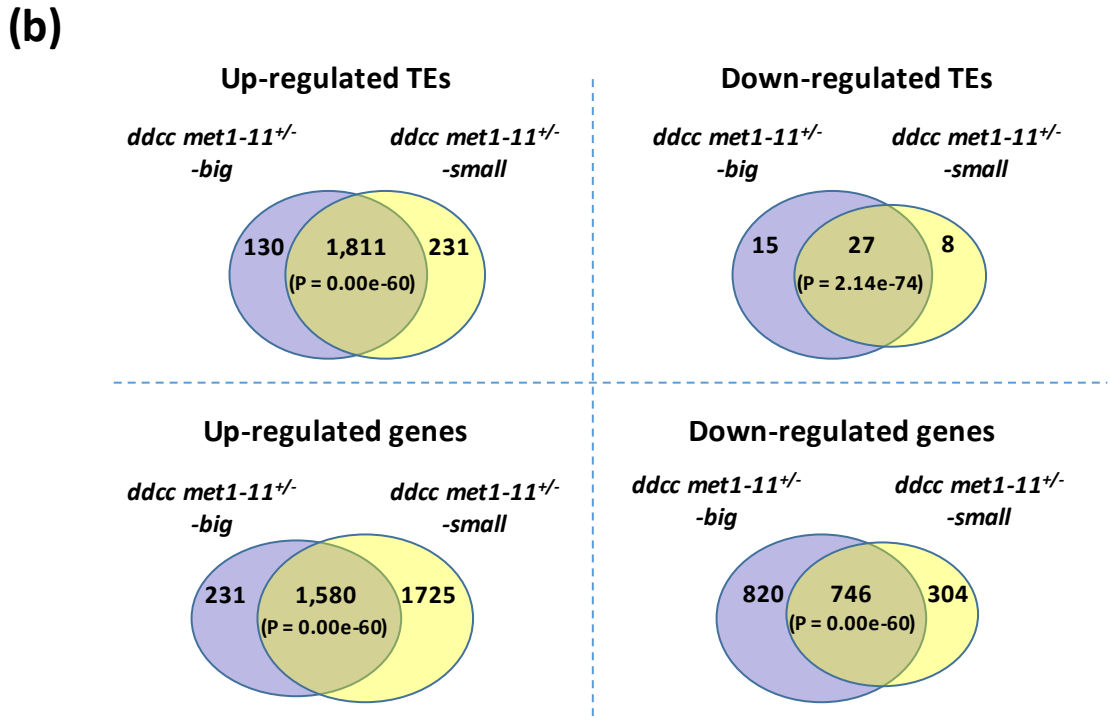


Fig. S6 The up-regulation of cell cycle-, DNA replication- and DNA repair-related genes in *Arabidopsis ddcc met1-11*$^{+/-}$ seedlings is not due to loss of DNA methylation in the promoters of these genes. Plots of CG methylation levels in the promoter regions (2kb upstream of transcription start sites) of cell cycle-, DNA replication- and DNA repair-related genes in Col-0 and the big and small plants of *ddcc met1-11*$^{+/-}$.

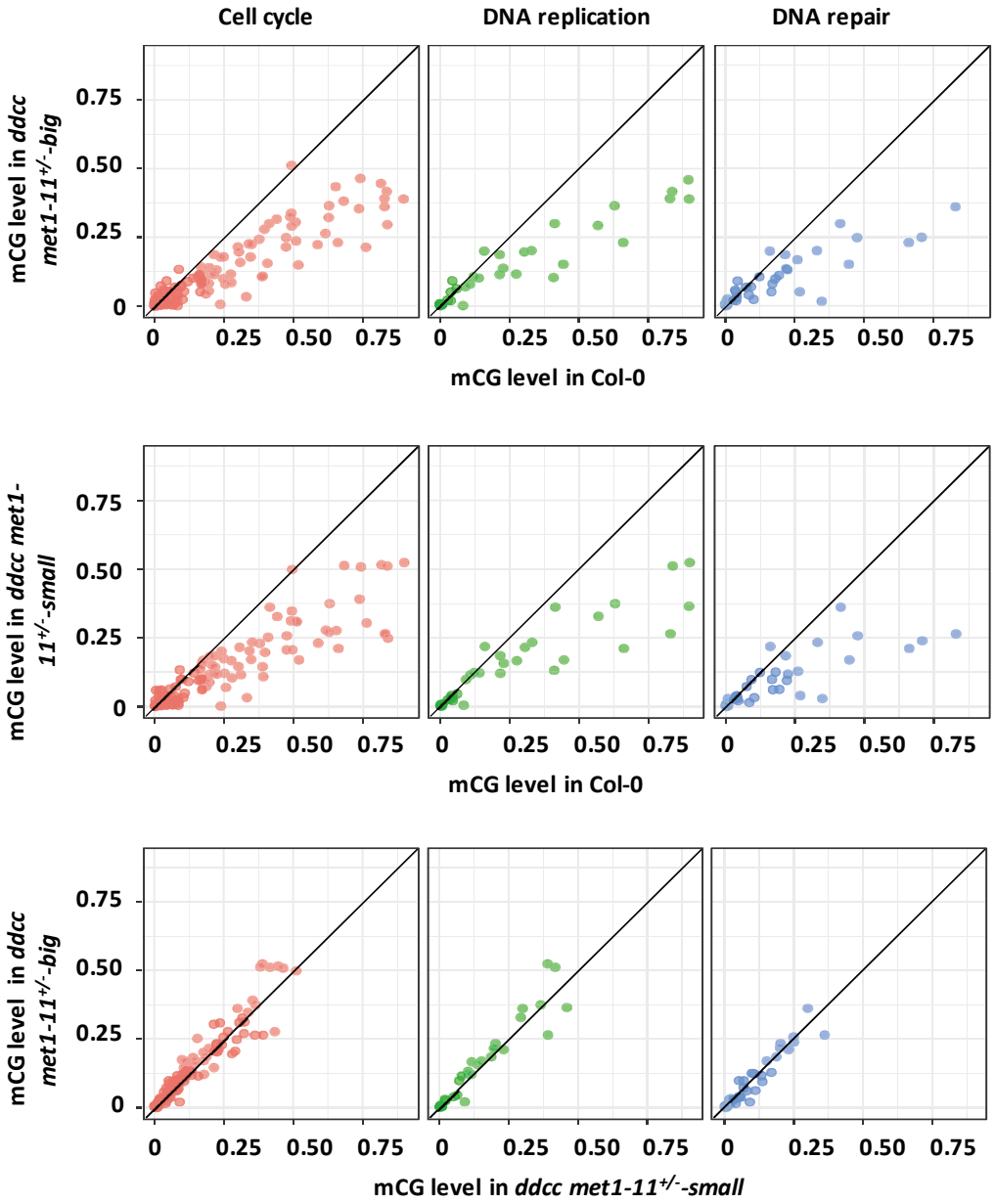
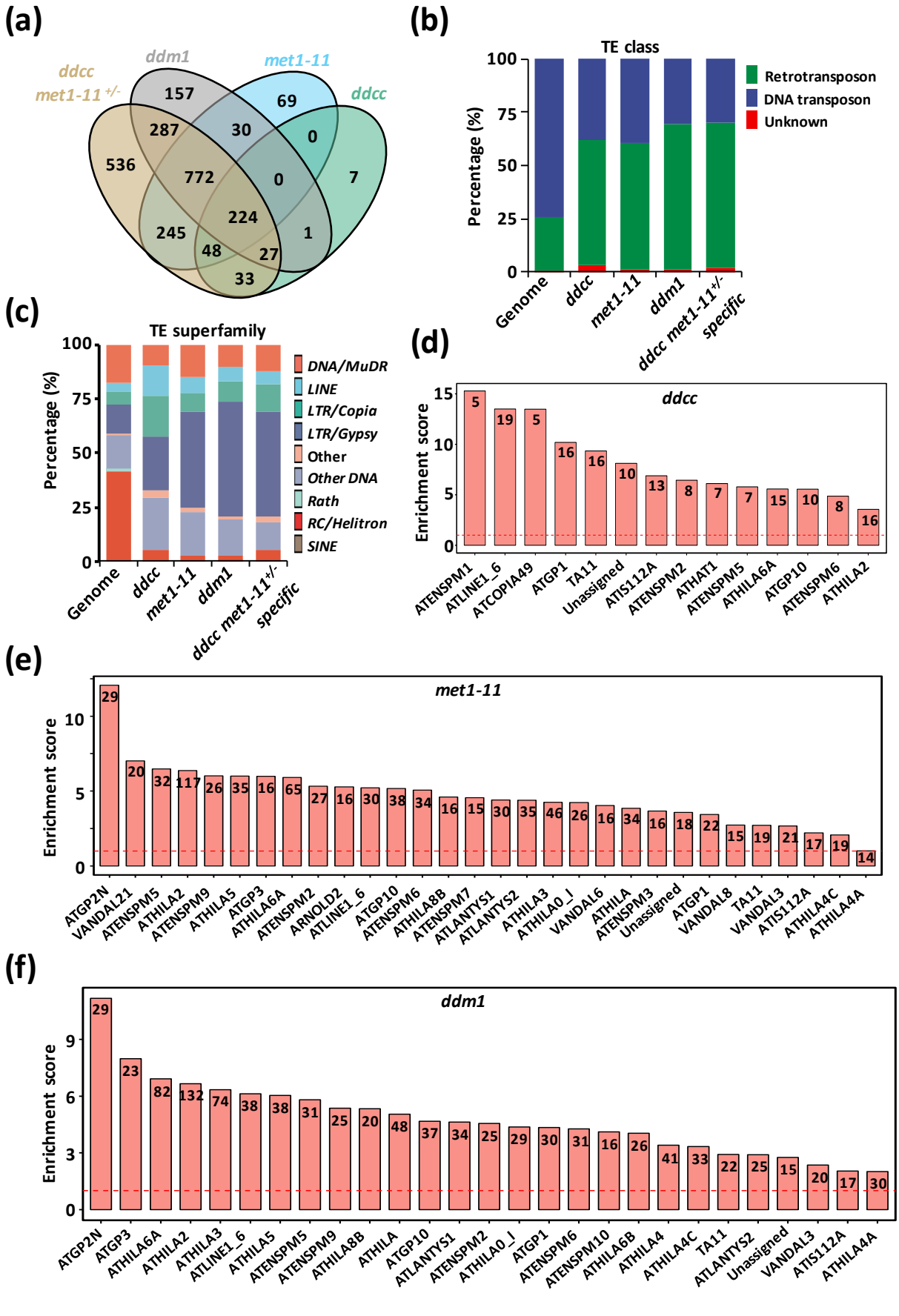


Fig. S7. The *Arabidopsis ddcc met1*$^{+/-}$ mutations specifically activate a group of TEs. (a) Overlap of up-regulated transposable elements (TEs) in *ddcc met1-11*$^{+/-}$ (activated TEs in the 'big' and 'small' plants combined), *met1-11, ddcc, and ddm1*. **(b)** Classification of the activated TEs in *ddcc, met1-11, ddm1* and TEs specifically activated in *ddcc met1-11*$^{+/-}$ (536 TEs), The *y* axis represents the percentage of individual categories. **(c)** Proportions of different transposon superfamilies in activated TEs in *ddcc, met1-11, ddm1* and TEs specifically activated in *ddcc met1-11*$^{+/-}$ (536 TEs). **(d)-(g)** Enrichment scores of TE families of activated TEs in *ddcc* (d), *met1-11* (e), *ddm1* (f), and TEs specifically activated in *ddcc met1-11*$^{+/-}$ (g). Each enrichment score was calculated as the ratio of the percentage of a certain TE family in activated TEs to that in all TEs. Numbers in the bars indicate the absolute numbers of activated TEs in the corresponding family. The red dotted line marks the point of overrepresentation (enrichment score >1). **(h)** Genome browser representation of RNA-seq and WGBS data for *AT1TE44535*, *AT2TE21060*, and *AT5TE38645*, showing transcript and methylation levels of these TEs in Col-0, *met1-11*, *ddcc*, and *ddcc met1-11*$^{+/-}$. *big*: *ddcc met1-11*$^{+/-}$*-big*; *small*: *ddcc met1-11*$^{+/-}$*-small.*


(a)
ddm1
met1-11
ddcc
met1-11 +/-
157
69
287
30
0
536
772
0
7
224
245
1
48
27
33
(b)
TE class
Percentage (%)
100
75
50
25
0
Retrotransposon
DNA transposon
Unknown
Genome
ddcc
met1-11
ddm1
ddcc met1-11+/-
specific
(c)
TE superfamily
Percentage (%)
DNA/MuDR
LINE
LTR/Copia
LTR/Gypsy
Other
Other DNA
Rath
RC/Helitron
SINE
(d)
ddcc
Enrichment score
5 19 5 16 16 10 13 8 7 7 15 10 8 16
ATENSPM1 ATLINE1_6 ATCOPIA49 ATGP1 TA11 Unassigned ATIS12A ATENSPM2 ATHAT1 ATENSPM5 ATHILA6A ATGP10 ATENSPM6 ATHILA2
(e)
met1-11
Enrichment score
29 20 32 117 26 35 16 65 27 16 30 38 34 16 15 30 35 46 26 16 34 16 18 22 15 19 21 17 19 14
ATGP2N VANDAL21 ATENSPM5 ATHILA2 ATENSPM9 ATHILA5 ATGP3 ATHILA6A ATENSPM2 ARNOLD2 ATLINE1_6 ATGP10 ATENSPM6 ATHILA8B ATENSPM7 ATLANTYS1 ATLANTYS2 ATHILA3 ATHILA0_I VANDAL6 ATHILA ATENSPM3 Unassigned ATGP1 VANDAL8 TA11 VANDAL3 ATIS112A ATHILA4C ATHILA4A
(f)
ddm1
Enrichment score
29 23 82 132 74 38 38 31 25 20 48 37 34 25 29 30 31 16 26 41 33 22 25 15 20 17 30
ATGP2N ATGP3 ATHILA6A ATHILA2 ATHILA3 ATLINE1_6 ATHILA5 ATENSPM5 ATENSPM9 ATHILA8B ATHILA ATGP10 ATLANTYS1 ATENSPM2 ATHILA0_I ATGP1 ATENSPM6 ATENSPM10 ATHILA6B ATHILA4 ATHILA4C TA11 ATLANTYS2 Unassigned VANDAL3 ATIS112A ATHILA4A

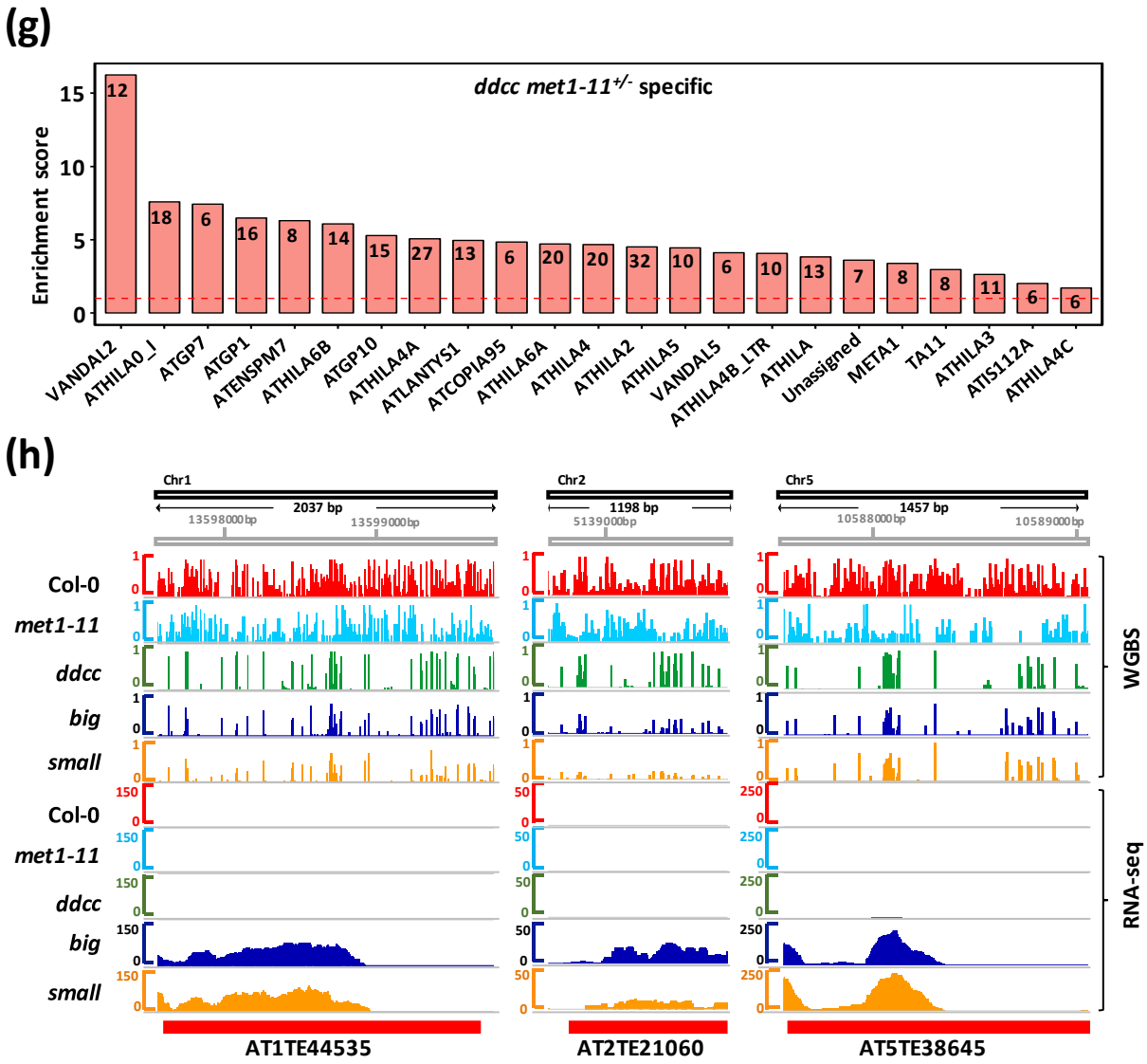

(g)
ddcc met1-11+/- specific
Enrichment score
0
5
10
15
12
18
6
16
8
14
15
27
13
6
20
20
32
10
6
10
13
7
8
8
11
6
6
VANDAL2
ATHILA0_I
ATGP7
ATGP1
ATENSPM7
ATHILA6B
ATGP10
ATHILA4A
ATLANTYS1
ATCOPIA95
ATHILA6A
ATHILA4
ATHILA2
ATHILA5
VANDAL5
ATHILA4B_LTR
ATHILA
Unassigned
META1
TA11
ATHILA3
ATIS12A
ATHILA4C
(h)
Chr1
2037 bp
13598000bp
13599000bp
Chr2
1198 bp
5139000bp
Chr5
1457 bp
10588000bp
10589000bp
Col-0
met1-11
ddcc
big
small
WGBS
RNA-seq
AT1TE44535
AT2TE21060
AT5TE38645

Fig. S8 CG and non-CG DNA methylation synergistically regulate chromocenter condensation in *Arabidopsis*.

(a) DAPI (4′**,** 6-diamidino-2-phenylindole) staining patterns of different types of nuclei. Type I nuclei have six to ten conspicuous, round chromocenters. Type II nuclei have elongated chromocenters and irregularly shaped regions with strong DAPI staining. Type III nuclei have few and small chromocenters. Bars, 2 μm.

(b) Percentages of three types of nuclei in the indicated plants. The error bars represent the SD of three independent experiments. Different letters indicate significant differences ($P < 0.05$, Fisher's Exact Test).

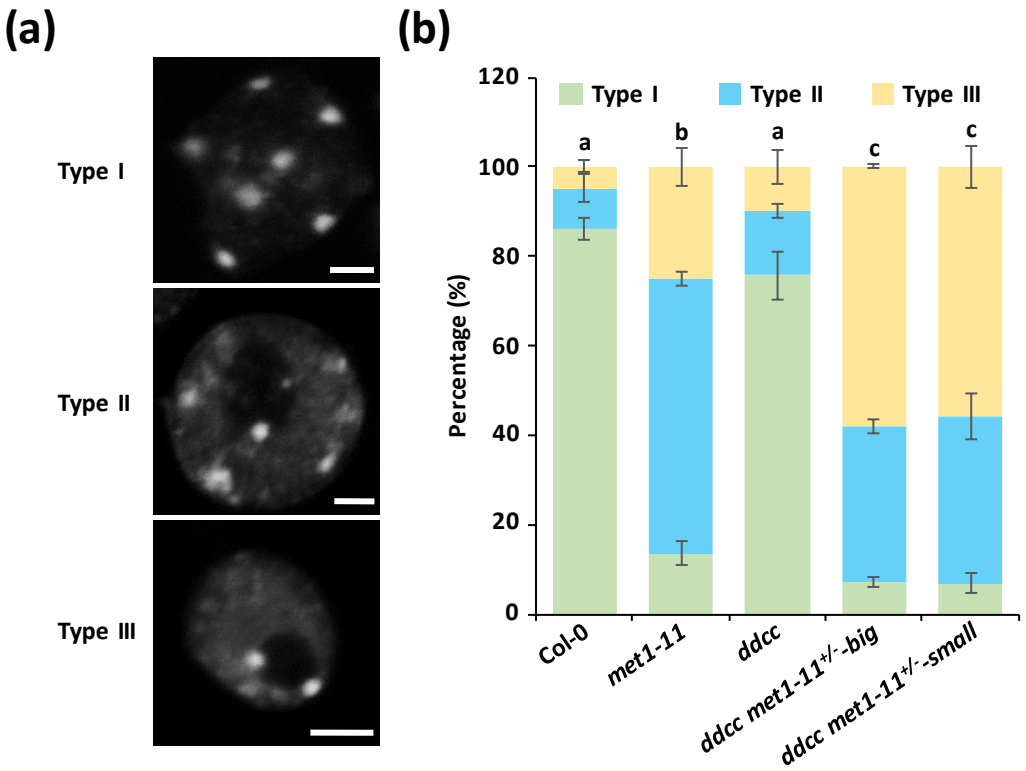


Table S1. Primers and sgRNAs used in this study.

Primer name	Sequence (5'-3')	Purpose
MET1-10-sg-F	GATTCGCGTTACGGTTAACGGCTC	CRISPR
MET1-10-sg-R	AAACGAGCCGTTAACCGTAACGCG	
MET1-11-sg-F	ATTGATTAGCAGCTAAACTAACTG	
MET1-11-sg-R	AAACCAGTTAGTTT AGCTGCTAAT	
Cas9-F1	GCTGTTTCTGCTCATTATCCTC	
Cas9-R1	ATTTCCAGTTTTACAAAGTGCG	
U6-26p-F	TGTCCCAGGATTAGAATGATTAGGC	
U6-29p-R	AGCCCTCTTCTTTCGATCCATCAAC	
U6-29p-F	TTAATCCAAACTACTGCAGCCTGAC	
ACTIN2-F	AACTCTCCCGCTATGTATGTCG	Real-time PCR in pollen
ACTIN2-R	AACCCTCGTAGATTGGCACA	
ANX1-F	GGTTTTAATCGGTGCCTTGTG	
ANX1-R	TCCCTGAGATGGTTGATTTGG	
ANX2-F	CTTCTGTGAATTCCAGCTTGC	
ANX2-R	GATACCTTTCCCACCTGTCC	
CDKA;1-F	GCCAAAAGCCCTTATTTCCTG	
CDKA;1-R	GTCCGTTGGTTTCCATTTAGG	
DAA1-F	GCGGATAAGATGGAGAAGGAG	
DAA1-R	GCATCGAAGTGCTTCAAATGG	
DAU2-F	TGGAGAAAACAGAGGAAAGCG	
DAU2-R	TTTGGTGAACGAGAAGGTGG	
DAZ2-F	CCGTTGCTTCATCCTCTAGTTC	
DAZ2-R	CCAAGACCAAAACTTTCTGCC	
DUO1-F	CTTCTCATCGACTCAAGGGC	
DUO1-R	ACCTCTTCCTCAACCAAACC	
DUO3-F	TGAGTTTGAGTGTGAAGAGATGG	
DUO3-R	TCCTTGCCTAGTTCTTTGCC	
FAS1-F	GCGCAAAACTCTGGACTTTC	
FAS1-R	TGACTTTATGGTGCCTGATGG	
FAS2-F	TCGAAGATTGTCATGGTCACC	
FAS2-R	AGTGCAGGTCTTGAAAGGTC	
FBL17-F	CTGCATAAGTTCAAGTTCAAGTCC	
FBL17-R	CCAAGTCCACATAAAACACCAG	
HAP2-F	GTTTGATTGCTCCAAAGGCG	
HAP2-R	GAGTCCTTGAGTATAGCTGTGC	
MGH3-F	AACTGCGAGGAAATCTACTGG	
MGH3-R	TTTGCGGATCTCACGAAGAG	
MSI1-F	GGAGTAAGTTCAAGCAGGGTC	

MSI1-R	ACATCTTCCACAACCCCTTC	
MYB97-F	ATCATACAACTCCACTCTCAGC	
MYB97-R	CTCGTTATCTGTTCTGCCTGG	
MYB120-F	GTAACAAATGGGCTCGCATG	
MYB120-R	GGATGGAGTTGATGGTTAGGG	
PCR11-F	AGCCCAATACAATCTCAAGGAG	
PCR11-R	TCGTTCCATATTCCCATGCC	
RBOHH-F	AGCTTACGGTTCCTTGATCG	
RBOHH-R	ACTCCATTGTGTCTCCCATTC	
TES-F	CAGTCTCAGGGAAATAGCAGG	
TES-R	ACTACGGTTAATGTGGCTTCC	
UBQ10 –F	CCAAGATCCAAGACAAAGAGGG	
UBQ10–R	TGAGAACAAGATGAAGGGTGG	
WIP1-F	TTAACAGAGGCCATCAACGG	
WIP1-R	TGAATTTGCTTCGCTGACTTG	
WIT1-F	AACCGCCAAAGATATCGGG	
WIT1-R	CACCATCAACACACAATTCTCTC	
ATR-F	GGACTCCGCCATATTCTACAAG	Real-time PCR
ATR-R	GGTCTTCAGCATCTCATACTCC	
BRCA1-F	CCATGTATTTTGCAATGCGTG	
BRCA1-R	TGTGGAGCACCTCGAATCTCT	
CDKB1;1-F	TGGCATGTTTACCCTAAGTGG	
CDKB1;1-R	TTCGGCTGGATTGTACTTGAG	
CDKB2;1-F	AGAGTCTCCCTGAAAGATTGC	
CDKB2;1-R	GTGCCACCCATTACTCTACAG	
CYCD3;3-F	AGATTGGTATACAGCTTCGTCAC	
CYCD3;3-R	TGGAGTTCTTGAGTAGATGTGAATC	
E2F3-F	TTGGAAGGGAGTTGATGCG	
E2F3-R	GGGCGAGGTTTTCAATTTCTG	
MCM2-F	CAGTTACAAGGGAGTGGACG	
MCM2-R	ATGGCTTCATGGATACTCACC	
MYB3R-4-F	TGCTTCTGATTCGTGCTCATC	
MYB3R-4-R	TCTTGACTGAACTGTGCCAG	
ORC3-F	GGACTTCCTAACAGCCCAAA	
ORC3-R	AAGGAACACTCTCAACTGGC	
PCNA1-F	ACCGCTAACATTGTGCTCAG	
PCNA1-R	CTGGCTCCTTCATCTCTATCAC	
RPA70C-F	TTGGGTTCGGTTATAAGACTCAC	
RPA70C-R	AGCTATGCCCAGGTTCTTTAG	
TCP15-F	TTTGGCTTCTGGTTATGGAGG	
TCP15-R	GTCACGGTTTTGCTGGTTG	

TOP3A-F	GCACGACCACATAAAGAAACTC	
TOP3A-R	AGGGCTCTGAGATTTGGTTTC	
WEEI-F	AGAACACTGTTGGCGAAGAG	
WEE1-R	GTCCCTTCCATCTTCCGAATC	