Supplementary Materials

Contents

(Supplementary Table S1 is shown in a separate file.)

Supplementary Methods

Immunoprecipitation and Microarray Analysis of Methylated DNA.
The genomic DNA was sonicated to produce random fragments ranging in size from 100 to 1,000 bp. Immunoprecipitation (IP) using anti-methylcytosine antibody was performed as described by Weber et al (2005). DNA samples were sent to NimbleGen for labeling and hybridization. Two independent experiments (experiments A and B) were performed using different plants and different probe amplification methods. In the experiment A, the immunoprecipitated DNA was amplified following the NimbleGen's instruction using GenomePlex Complete Whole Genome Amplification Kit (SIGMA, WGA2). In the experiment B, the probe was amplified by linear amplification method described by Liu et al. (2003). Although the amplification methods are different, response to the *ibm1* mutation in each gene gave consistent results (Supplementary Fig. S1-S13). The analyses in Figure 1, 3, 4, 5A, 5C, and 5D are based on results of experiment A.

Data Normalization. The Arabidopsis genome is covered by three 385K chips of NimbleGen. The first chip covers entire chromosome 1 and part of chromosome 2. The second chip covers entire chromosome 3 and parts of chromosome 2 and 4. The third chip covers entire chromosome 5 and part of chromosome 4. In order to integrate the results from three chips, the IP/input signal values of each chip was normalized using average of one chromosome for each chip: chromosome 1, 3, and 5 for chip 1, 2, and 3, respectively. Values for part of chromosomes were not used, in order to avoid bias caused by the different proportion of methylated pericentromeric regions and arm regions in the part of chromosome. As methylation level of transposons are similar between *ibm1* and wild type (Fig. 1C, 2N-2P), we used value of y-intercept of Figure 1C for comparison between *ibm1* and wild type.

Relationship between DNA Methylation and Gene Expression. Data used in the analyses of gene expression levels and tissue specificities were from a previous study analyzing the genome-wide expression pattern in various developmental stages and tissues of Arabidopsis (Schmid et al. 2005). For Fig.

4A and 4B, sum of expression in all of stages and tissues was used. For tissue specificity (Fig. 4C and 4D), entropy levels of each gene were calculated as described by Zhang et al. (2006). For Fig. 5C, 5D, and S15, expression profiles in wild-type leaf in our microarray analysis (data not shown) were used.

References for Supplementary Methods

Liu CL, Schreiber, SL and Bernstein BE (2003) Development and validation of a T7 based linear amplification for genomic DNA. *BMC Genomics* **4**: 19

Schmid M et al. (2005) A gene expression map of Arabidopsis thaliana development. *Nat. Genet.* **37**: 501

Weber M et al. (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**: 853-862

Zhang, X. et al. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. *Cell* **126**, 1189-1201

Supplementary Table S2

Genes responding to *ibm1* were not necessarily located near transposons.

	Class I	Class II	Class III	Total
Total	651	2,461	23,931	27,043
	(0.024)	(0.091)	(0.88)	
Transposon-flanked	156	841	7,813	8,710
	(0.018)	(0.097)	(0.90)	

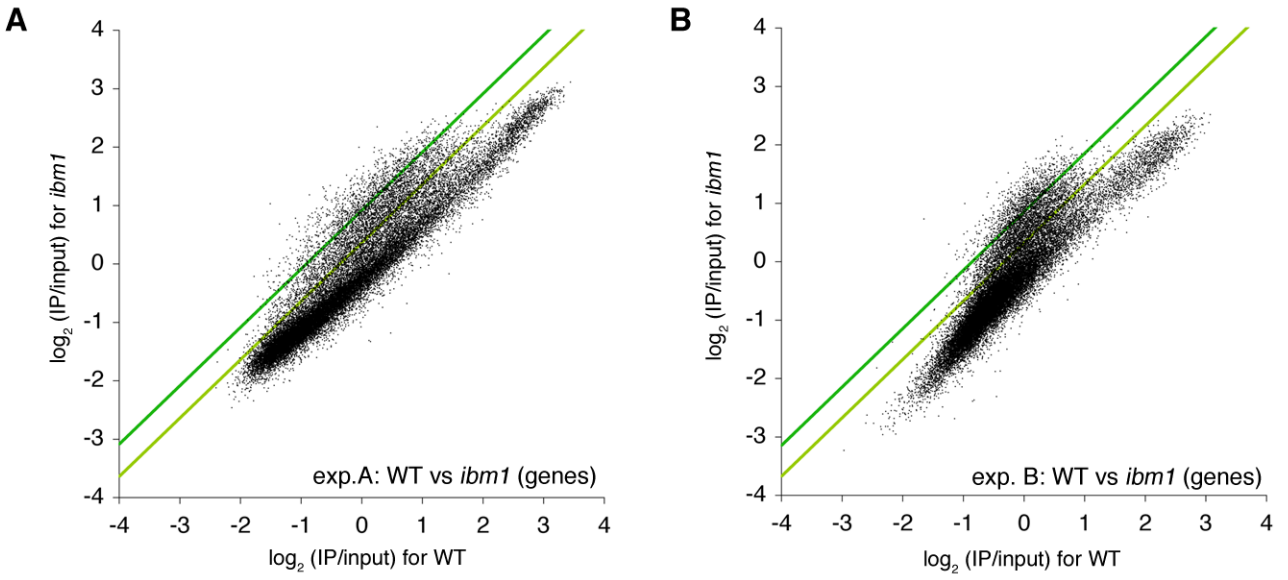
Genes flanked by transposons were counted for each Class of genes. We defined the “transposon-flanked” genes as those flanked by transposons within 1,000 bp, and that do not overlap with transposons more than 1,000 bp. Proportion of each Class within the total and transposon-flanked genes are shown in the parenthesis. Class I genes were not enriched in the transposon-flanked population.

Supplementary Table 3

Primer sequences and restriction enzymes used for the bisulfite sequencing.

Target	Primers	Restriction enzyme	Sequence (5' to 3')
AT2G16920	AT2G16920-F1	*Hin*dIII	TGAGAATGATTYTGGGAGGAA
	AT2G16920-R1		CCCCRACTATTACTRCCCTCA
	AT2G16920-F2	*Hin*dIII	AAGAAGGTTGATYAAGAYTGGAAAA
	AT2G16920-R2		TTCCRACCTRCTTATCATACCC
	AT2G16920-F3	*Eco*T14I	GGAYAGGYAGGGGAAATGAAGT
	AT2G16920-R3		CCCAAAACATRACCAACCAA
AT4G00450	AT4G00450-F1	*Eco*RV	TTGTAATTGAGGTTYTGAAYYTTTGG
	AT4G00450-R1		TCACTTTTATAAACCAARTARCTCTCAATA
	AT4G00450-F2	*Eco*RV	GGAGGAGAYAAAGAAAAGATGGA
	AT4G00450-R2		CACCATATACCACCATTTAAAAT
	AT4G00450-F3	*Eco*RV	AATGGAYAAAAGATGTTATTGAATATTTG
	AT4G00450-R3		AACAATACTCTCTAAAACACCATACACAAT
AT4G16310	AT4G16310-F1	*Ssp*I	GAAGAATTTAGTTATGTGTGATTGTGG
	AT4G16310-R1		ACATCTTAAARTACACCCTTCTTATAA
	AT4G16310-F2	*Hin*dIII	AAGGAAATTGAAATTGGTAGAG
	AT4G16310-R2		AAACCARCACTCATTCTCTTTTC
	AT4G16310-F3	*Hin*dIII	TGTTATATTTGTGTGGTTYGAT
	AT4G16310-R3		CTATAACTTTCTTCTCATCTATAACCTC
AT4G31440	AT4G31440-F1	*Nde*I	GTTTTTGGGGAGATGAGATTG
	AT4G31440-R1		CAAAARRACCCAAAACCTTTCC
	AT4G31440-F2	*Nde*I	GGTGYTTGGGAAAGAAGATGG
	AT4G31440-R2		CACTCCRCTRAAACCCCTCC
	AT4G31440-F3	*Nde*I	TGAGGAAGAGGATGGAGAATATTG
	AT4G31440-R3		CCAAATTTCTTTTRATCTTACAACAACAC
At5G35210 Top-Strand	At5G35210-tF1	*Eco*T14I	AGATTTGGTYATTGGGTTAGGTTT
	At5G35210-tR1		ATCAAATCCAATTTCARATTCTTC
	At5G35210-tF2	*Dra*I	ATTGGGATGAAATTGAAGATT
	At5G35210-tR2		CTAATTCCAACAAATRCAAAACAAATT
	At5G35210-tF3	*Dra*I	GAGGAGYTGTAYAGTTTGGAATGGAT
	At5G35210-tR3		CCAATTTTTTCCCCTRTATTCCACAT
At5G35210 Bottom-Strand	At5G35210-bF1	*Nde*I&*Hin*dIII	TTTGTGTAATGATTATTGTATGTGTTGGG
	At5G35210-bR1		TTTTCRTTRCCTCCATTCATACATAATTT
	At5G35210-bF2	*Dra*I	AGGGGAATAGATTGAAAGAT
	At5G35210-bR2		AAATARAATACTACTCRTTACCCATT
	At5G35210-bF3	*Eco*T14I	GGTTTTAGYGAATYTGGGGTGGA
	At5G35210-bR3		CTTTATATCTRCCCATTTRAACTCAAT

At5G01580 Top-Strand	At5G01580-tF1	*Nde*I&*Hin*dIII	TTGTTTAATGAGAYAYAAGAATAAAG
	At5G01580-tR1		CTCAATCCAACCATTTCTAAACAATC
	At5G01580-tF2	*Nde*I&*Hin*dIII	TTTGAYTGTTAATGTGAATTATATGTG
	At5G01580-tR2		TTCTRTTCACCAAARACATATAAAT
	At5G01580-tF3	*Nde*I&*Hin*dIII	TAGATTTGTAYYATGGGTGGTTGTGAATAA
	At5G01580-tR3		TAACCTTTARCCTCCAAAATCTTCA
At5G01580 Bottom-Strand	At5G01580-bF1	*Dra*I	ATGTAATGGTATTTGGGATATTATAAA
	At5G01580-bR1		CCCAAGTTTRCTATTCCAATCACTAA
	At5G01580-bF2	*Nde*I&*Hin*dIII	GTGTAGAAYGAATTGGAGAGGAG
	At5G01580-bR2		AACATTARTTAACTTATACARCCACAACA
	At5G01580-bF3	*Nde*I&*Hin*dIII	TTGAAAGTTATTAGYTTAATTTATAGGA
	At5G01580-bR3		CTTAATCTCATCCATTRATCTCCA
AT3G16950 Top-Strand	AT3G19650-tF1	*Nde*I	AAAAGGTTAAATATGGGAAGGA
	AT3G19650-tR1		CTTTCTTRRATTTATCAAAACCCTCT
	AT3G19650-tF2	*Nde*I	GTAAATAGTAGAAAATAAGATTTTAAGG
	AT3G19650-tR2		CACAAATCTTCCCAATTCA
	AT3G19650-tF3	*Nde*I	TAAGGATAYTTTGGAGGTAA
	AT3G19650-tR3		CCCCTTTCCATCRATCACAC
AT3G16950 Bottom-Strand	AT3G16950-bF1	*Eco*RV	TAAAAGGTGGGATTTAGTGT
	AT3G16950-bR1		CAAATACTRCTTCTCTTTCTCTATACA
	AT3G16950-bF2	*Nde*I&*Hin*dIII	TGAATAAGAATAGATTTAAGTTTAA
	AT3G16950-bR2		CAATTARCAATRTTCTCCAATTTTCTA
	AT3G16950-bF3	*Dra*I	AATAYAGGAAGAGAYYAAAYAAAGAA
	AT3G16950-bR3		CTACCAAAATACRAAACAATCT
AT3G21215	AT3G21215-F1	*Eco*T14I	GGTATTGTTGGGGATTYAAATGYTTATGAT
	AT3G21215-R1		ATCAAACAATCARAAACTTTARCAAATTCA
	AT3G21215-F2	*Nde*I	GGTTGTGTTGTGGAAYAATAGTTGTT
	AT3G21215-R2		TATRCCTTTCTTRTCTCARAATCTTCAT
	AT3G21215-F3	*Eco*T14I	GAAGAAGAATTGAGGAGYYTGTTGAGTG
	AT3G21215-R3		CCTTCCTTTCTCTTCCCATATRRATTCTTT
AT4G35270	AT4G35270-F1	*Nde*I	TGYTTYAAGGGGAAGGYATTGTAGGAA
	AT4G35270-R1		TTCTCTRCAAAAACAACTTCCTCTCTAAC
	AT4G35270-F2	*Nde*I	GTTAGAGAGGAAGTTGTTTTTGYAGAGAA
	AT4G35270-R2		TCTTTTCTRTTTTTRTTCTTCTCTTCTCAC
AtMu1	ATMU1-bF1	*Eco*RV	GTGTTTATGATTATATAATTGTGTTATAAT
	ATMU1-bR1		CCTTCTCTTTCATTCARATTTTAATTTT
Ta3	JP1615[*1]	*Dra*I	GGTTTAATGTTGGTTTAGTGATATTYGGTTTAGT
	JP1616[*1]		AATCAAAAAACRAATAAACCTCRCTCTRATACCACTTATT
CACTA2	CAC2-bsF2	*Dra*I	CATATAAACCCCAAAATCAAATC
	CAC2-bsR2		ATGGAAAAGGAGAAGGAGGTAT



C

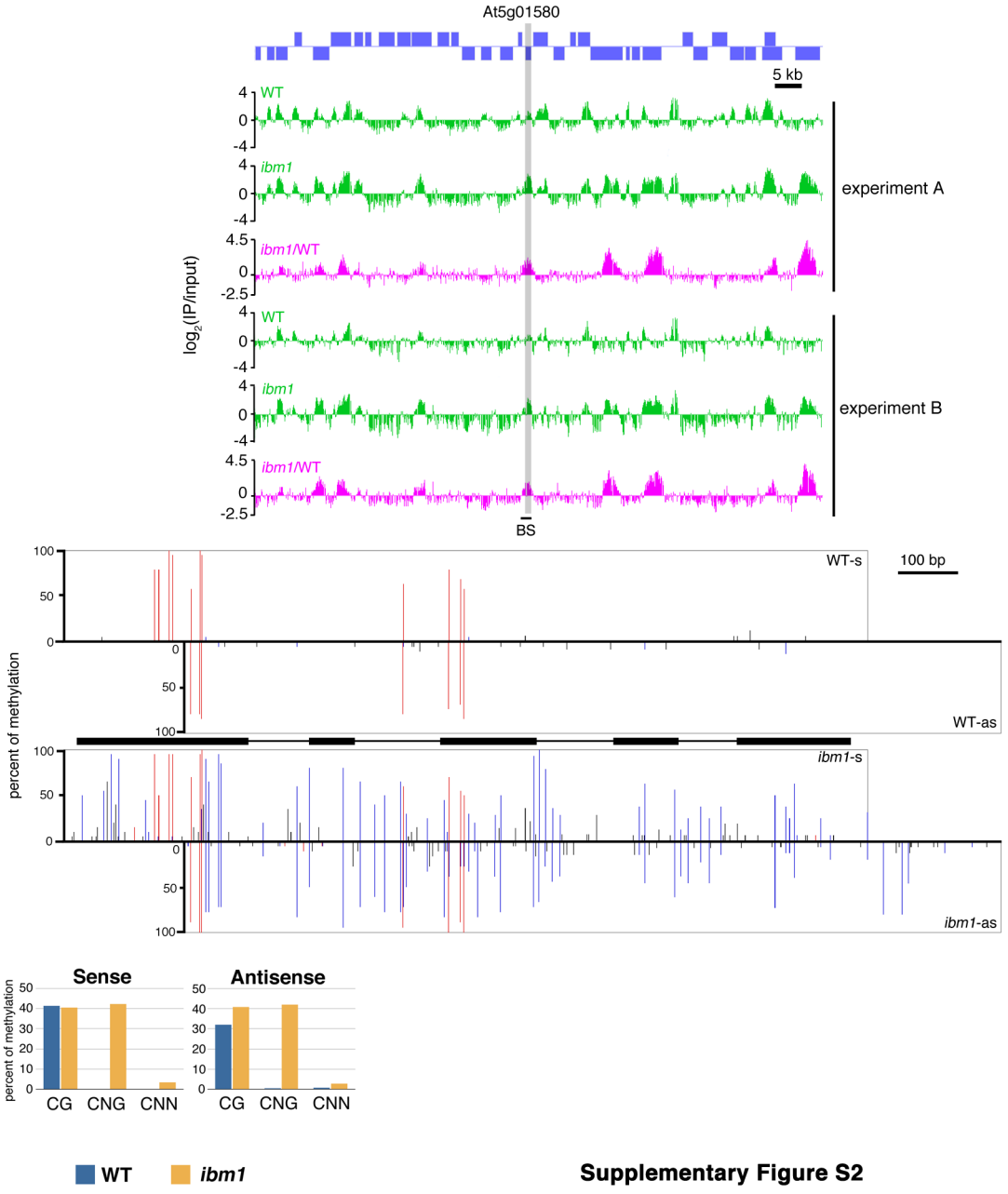
exp. B \ exp. A	1-1000	1001- 4000	4001-27043
1-1000	651	334	15
1001- 4000	296	1831	873
4001-27043	53	835	22155

Class I
Class II
Class III

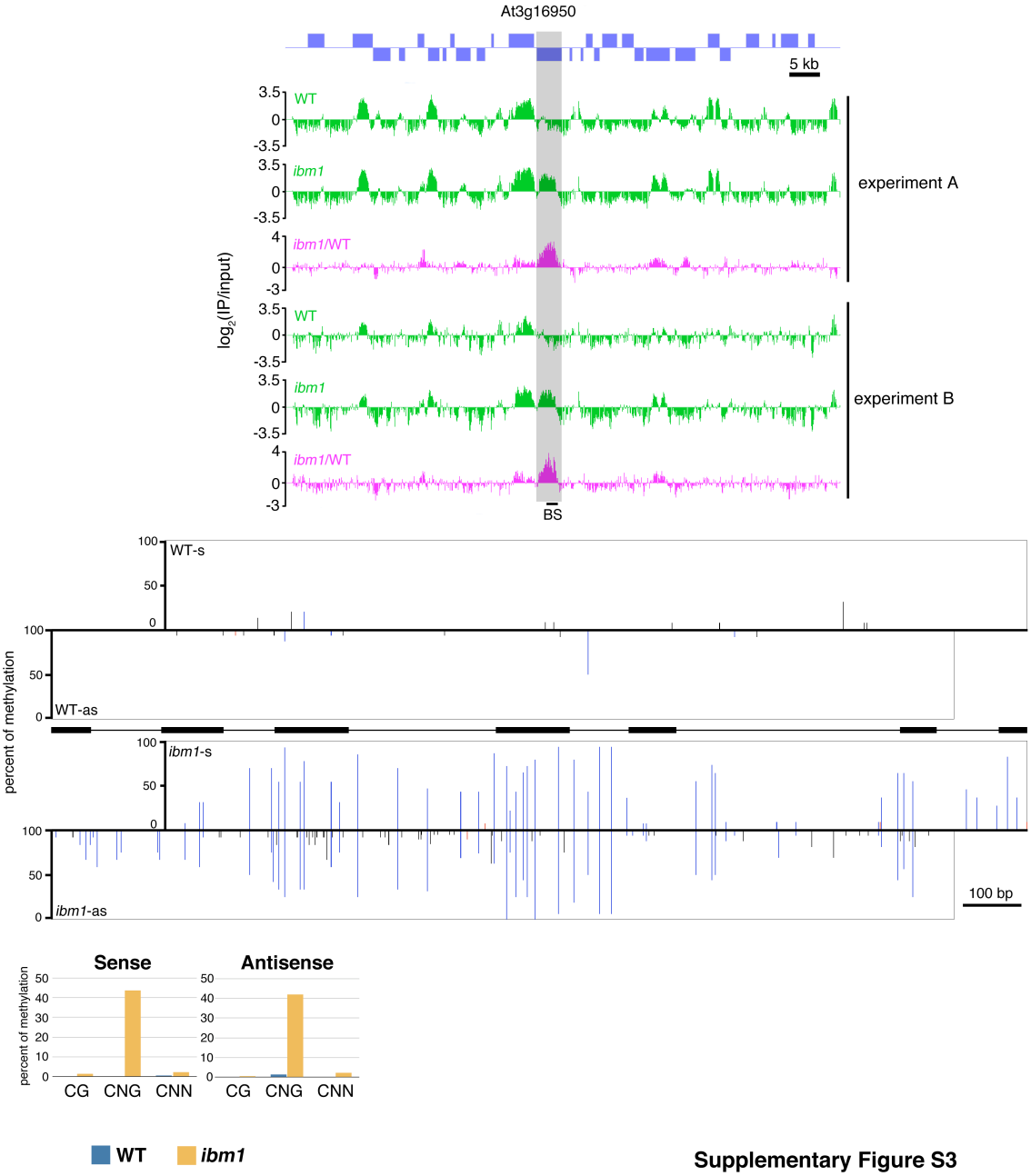
Supplementary Figure S1 Genes showing strong response to the *ibm1* mutation. (**A, B**) Each gene was compared between *ibm1* and wild type for the average of methylation signal of every probes covered by that gene. Experiment A (panel **A**) and B (panel **B**) were independent experiments using different *ibm1* and wild type plants and different probe amplification methods (See Supplementary Methods). (**C**) As at least 4,000 genes show significant changes, they were compared between experiments A and B. Among the 27,043 genes, we selected 3,112 genes (651 species of Class I genes, and 2,461 Class II genes) that reproducibly show most dense increase in methylation signal induced by the *ibm1* mutation.

Legend for Supplementary Figure S2-S13.

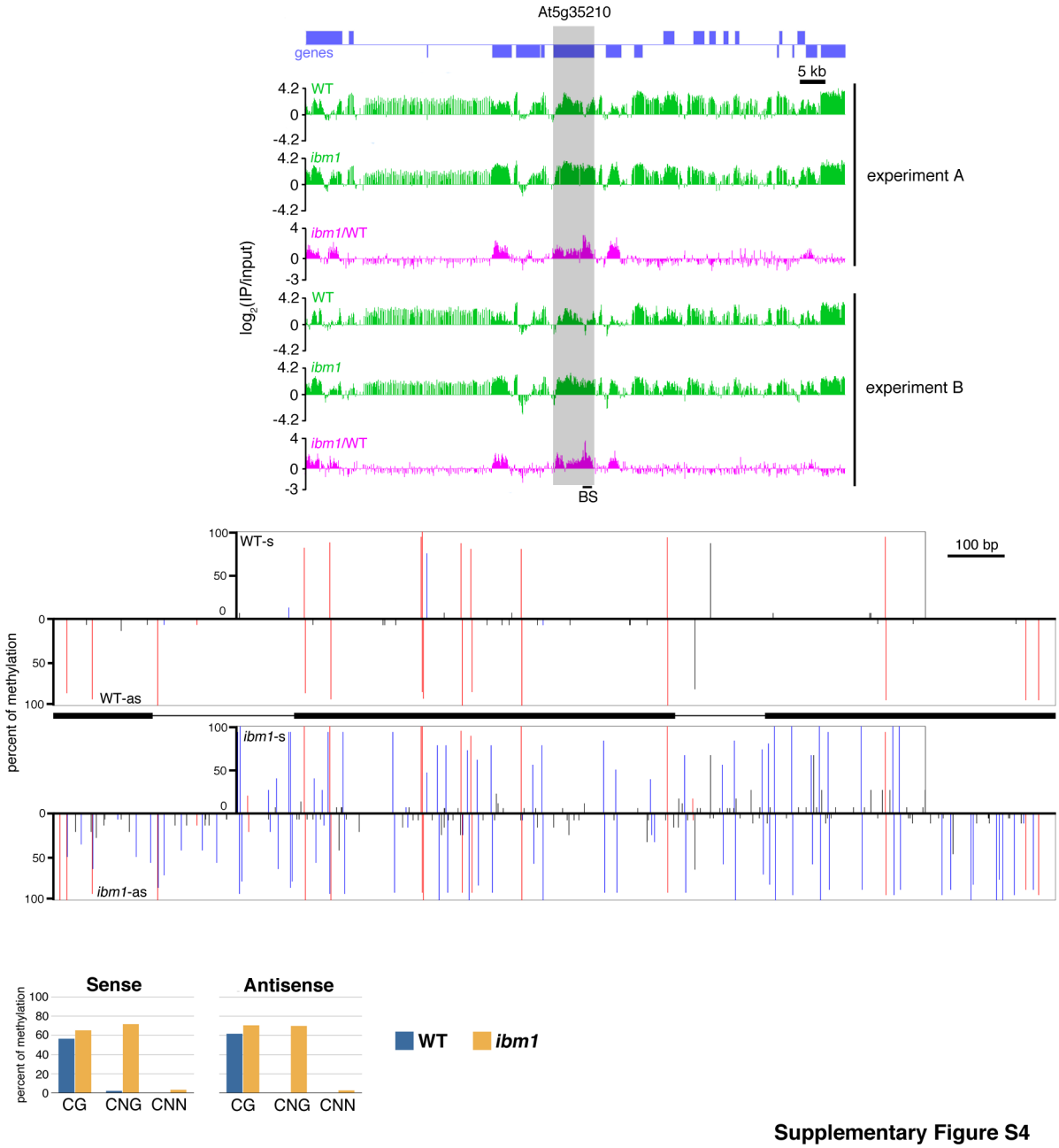
Full results for DNA methylation status in the 12 representative loci. Results are shown in the format of Figure 2 in the main text. Tiling array results of both experiments A and B are shown. In the Supplementary Figure S2-S4, bisulfite sequencing results of both sense and antisense strands are shown.



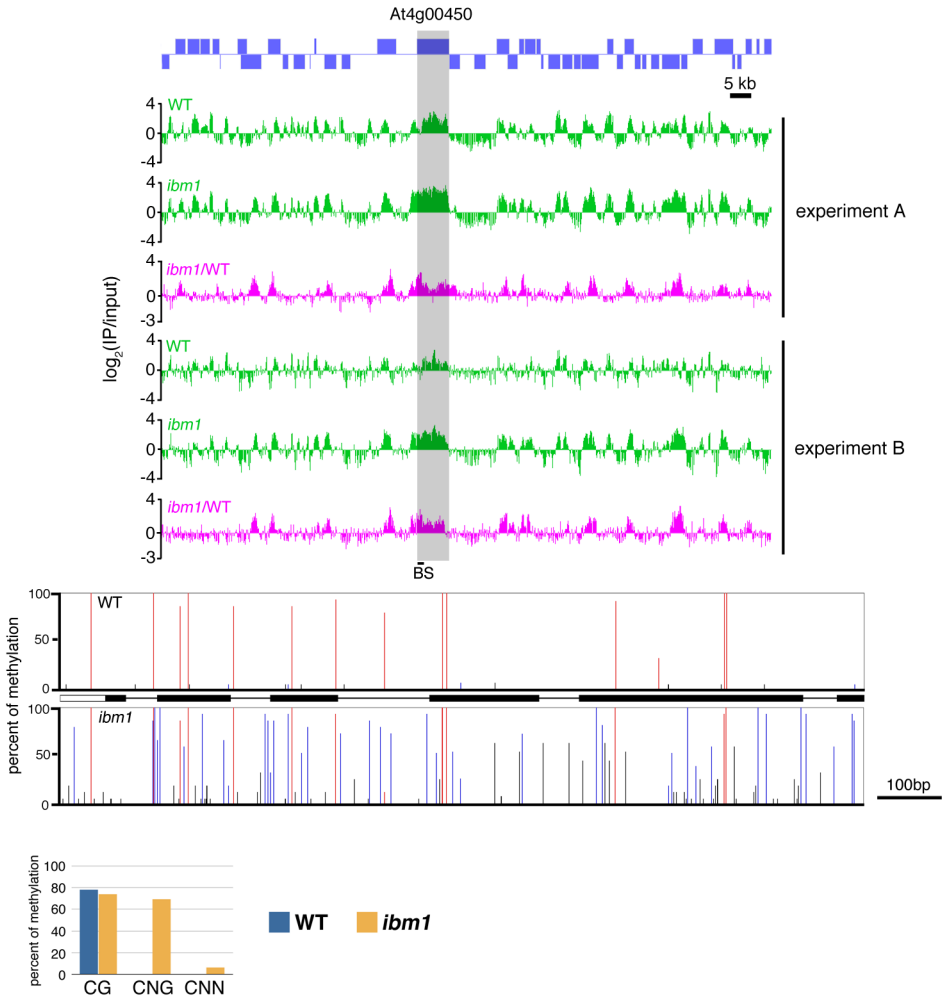
Supplementary Figure S2



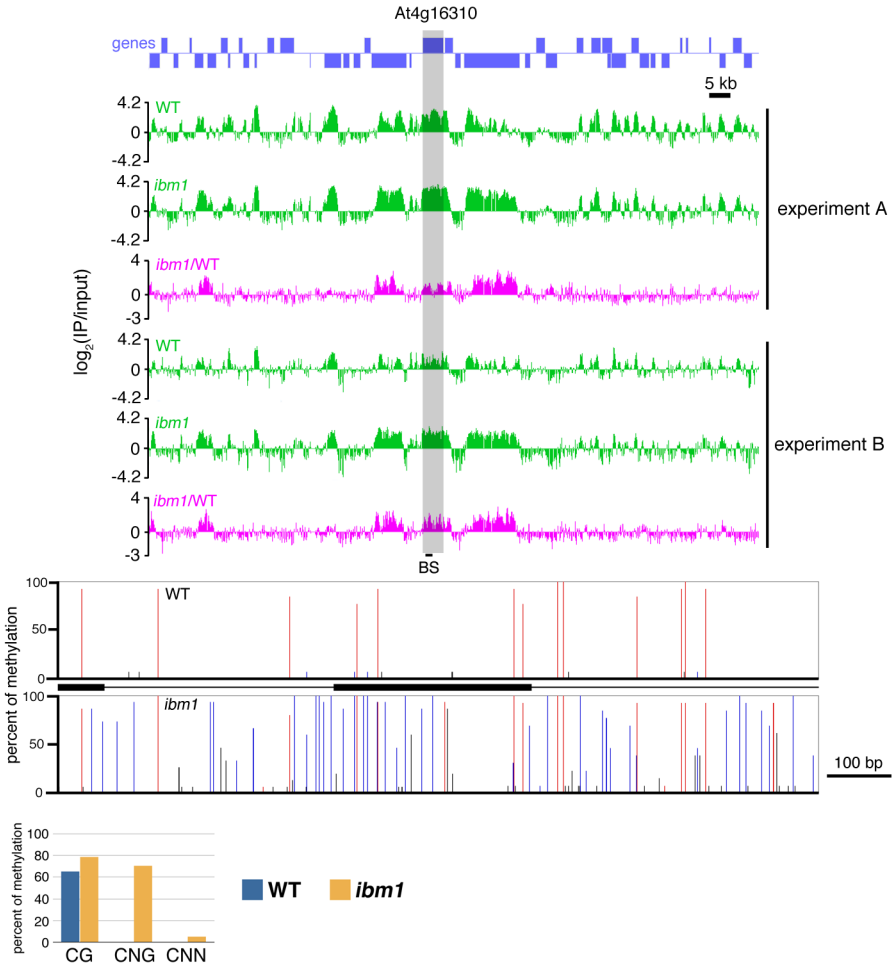
Supplementary Figure S3



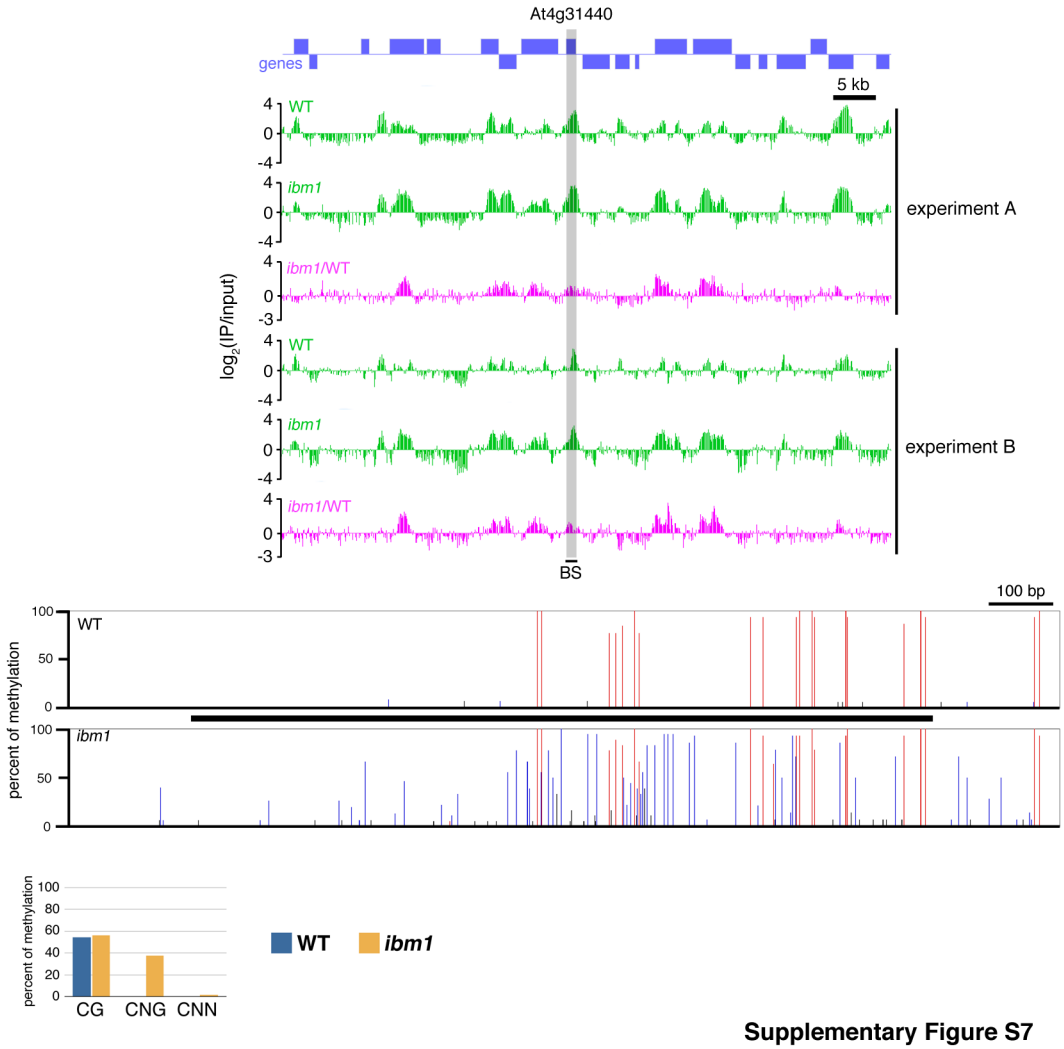
Supplementary Figure S4



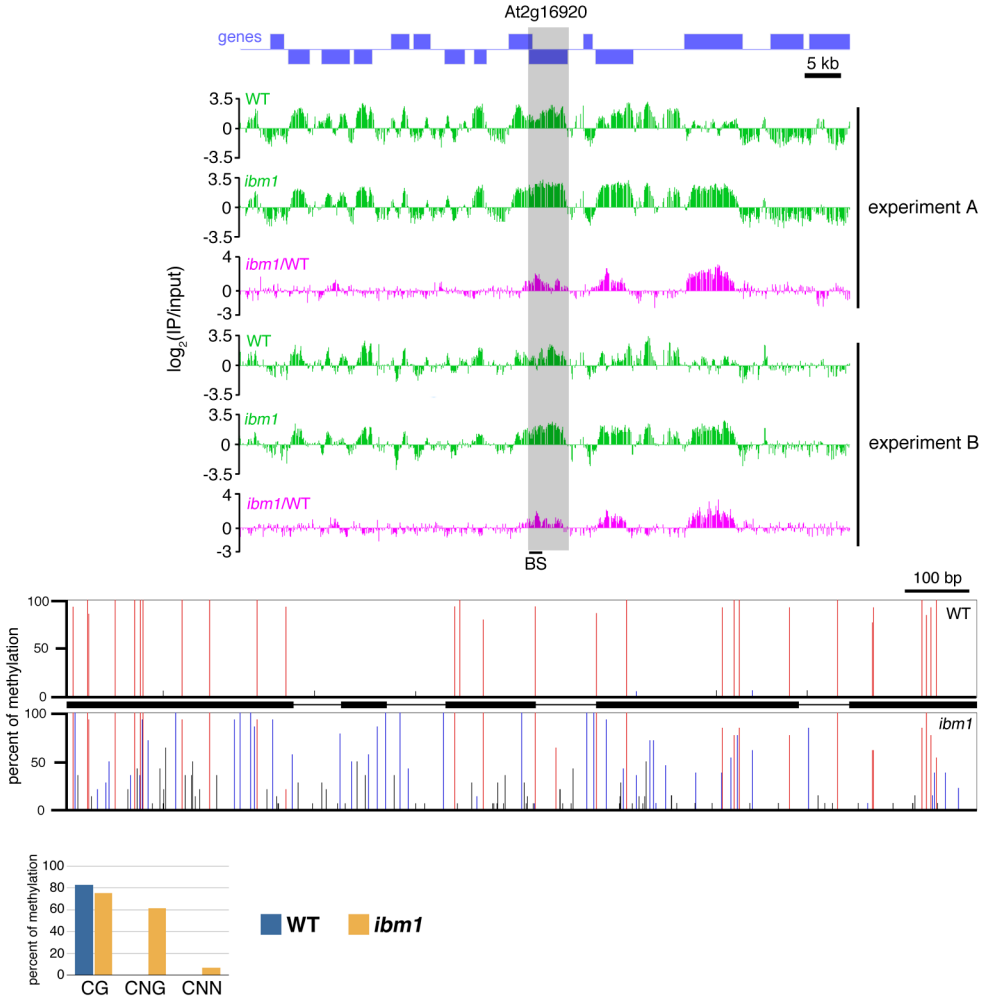
Supplementary Figure S5



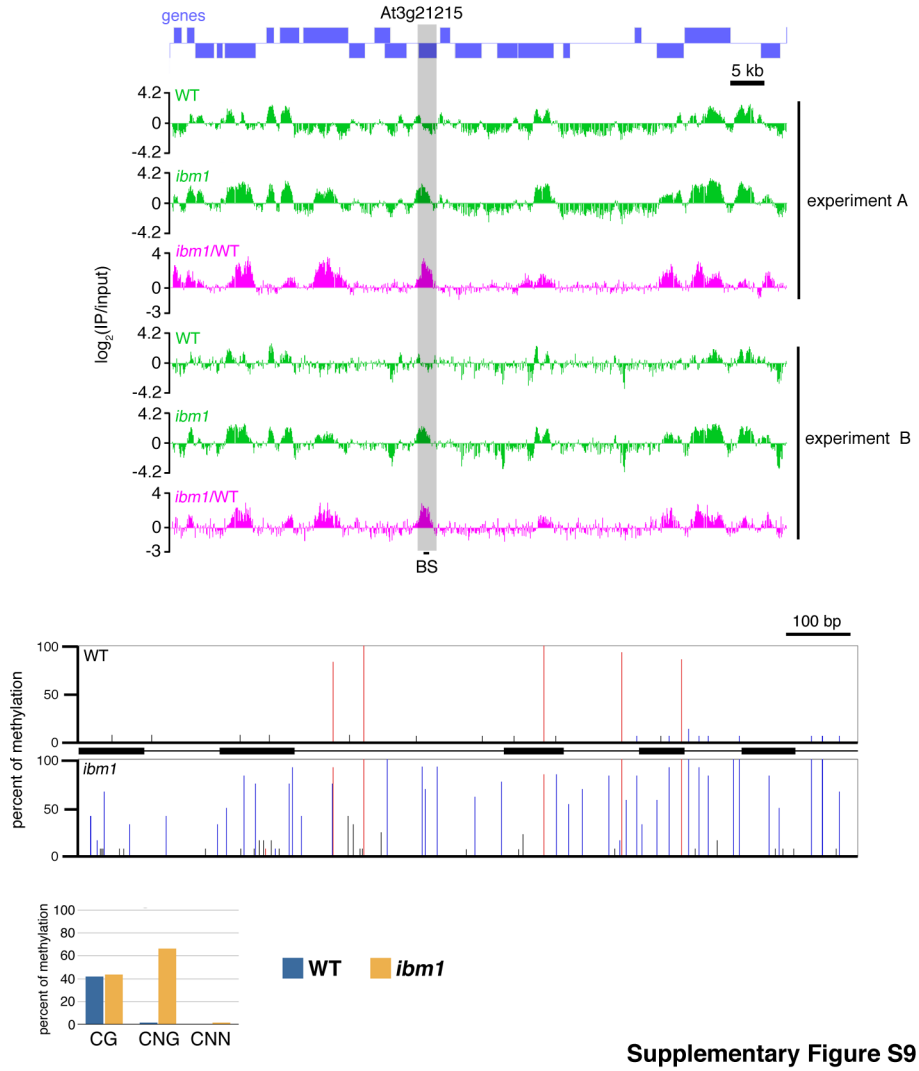
Supplementary Figure S6



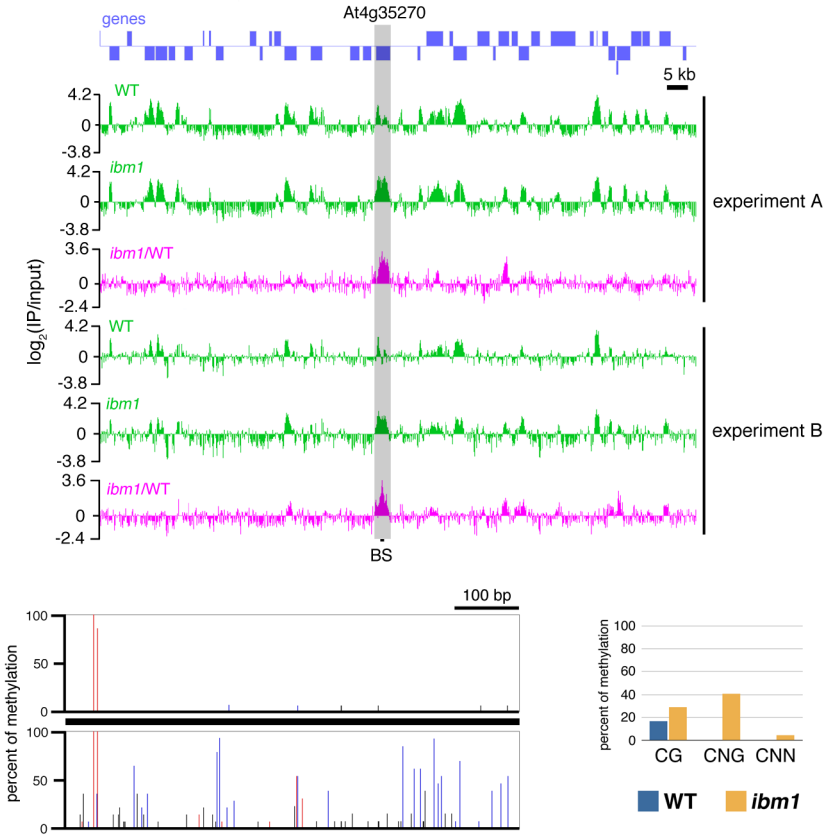
Supplementary Figure S7



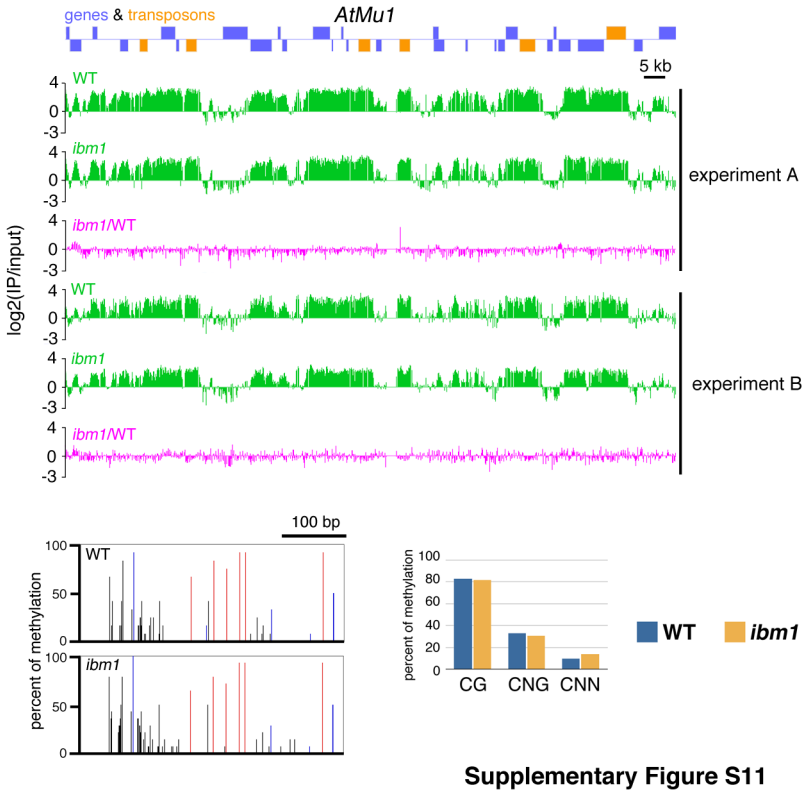
Supplementary Figure S8



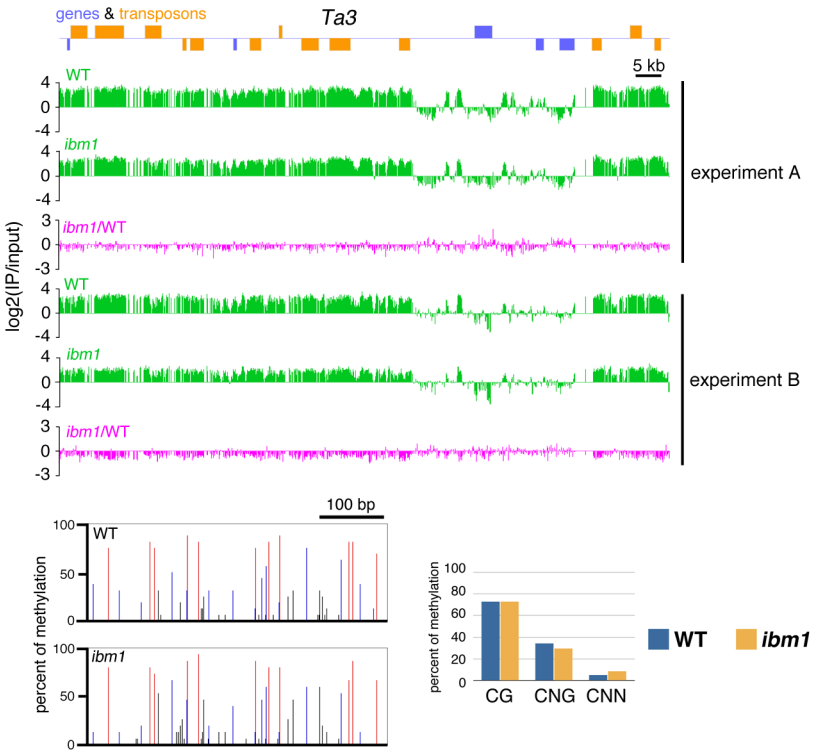
Supplementary Figure S9



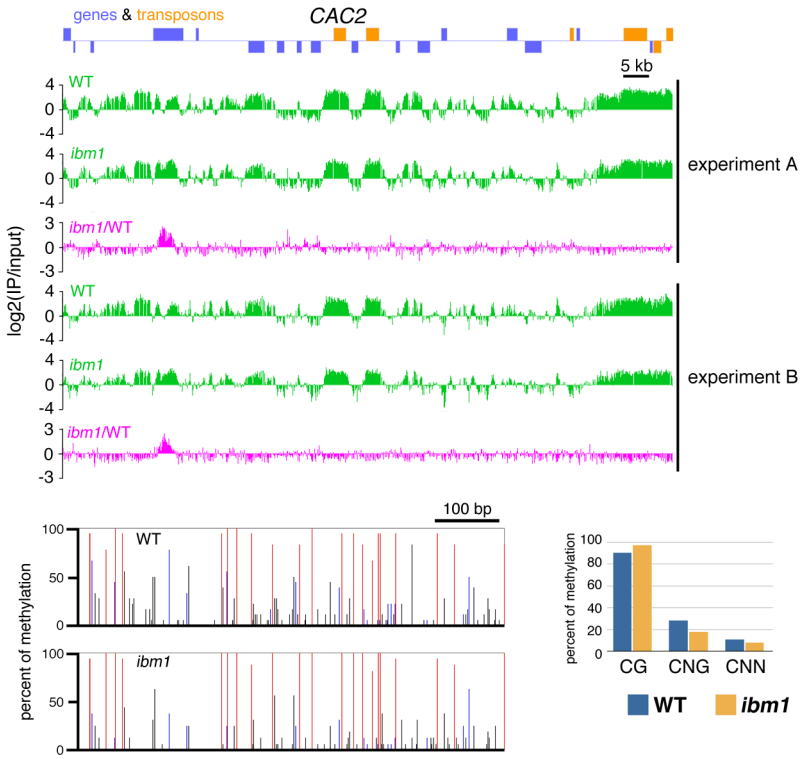
Supplementary Figure S10



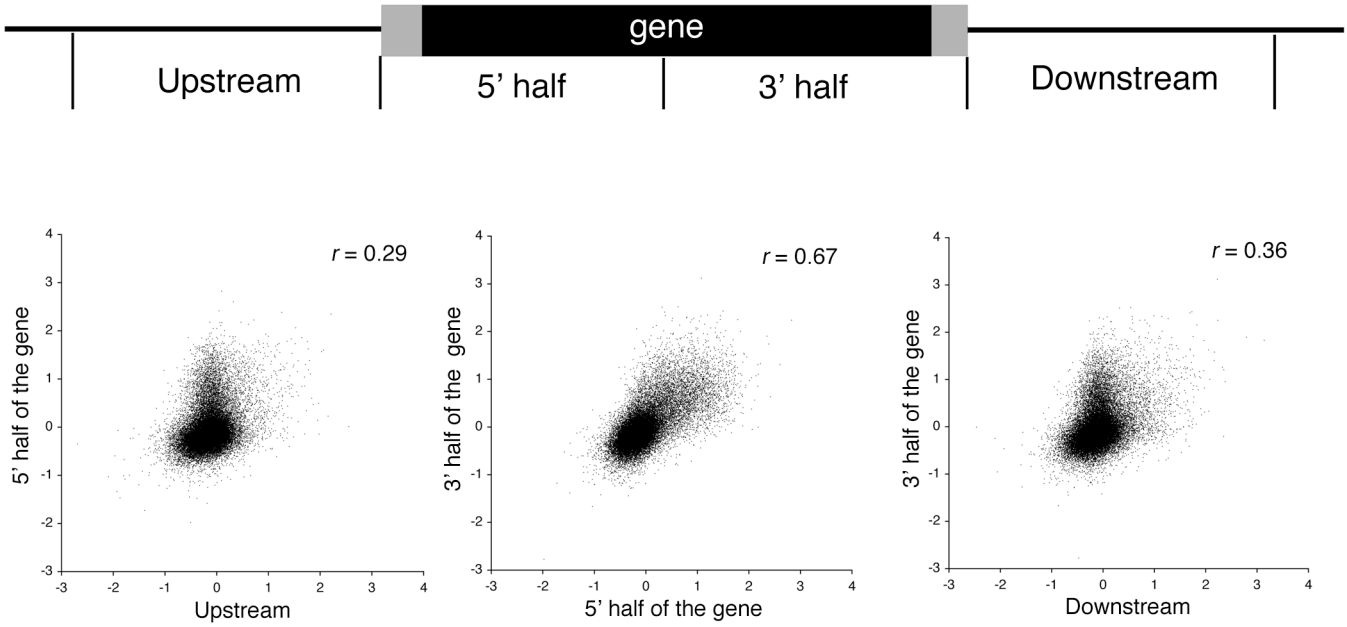
Supplementary Figure S11



Supplementary Figure S12



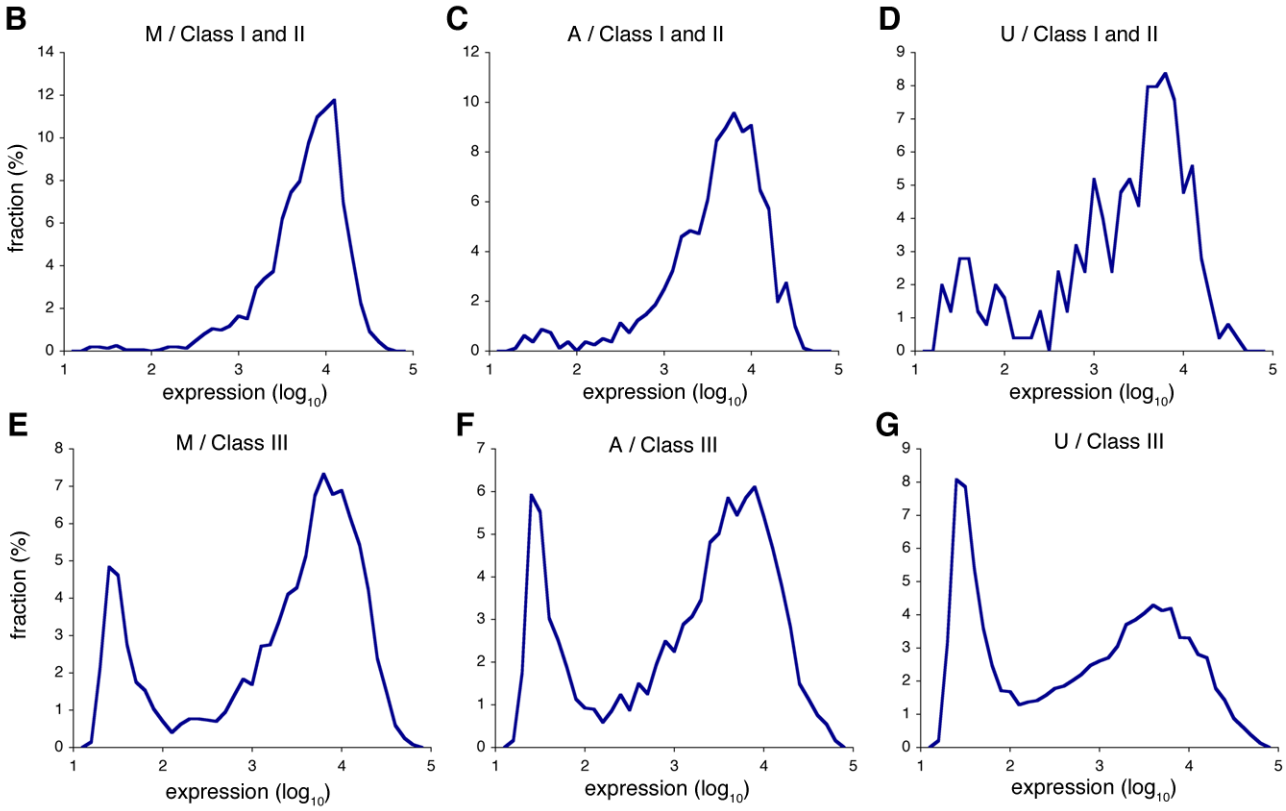
Supplementary Figure S13



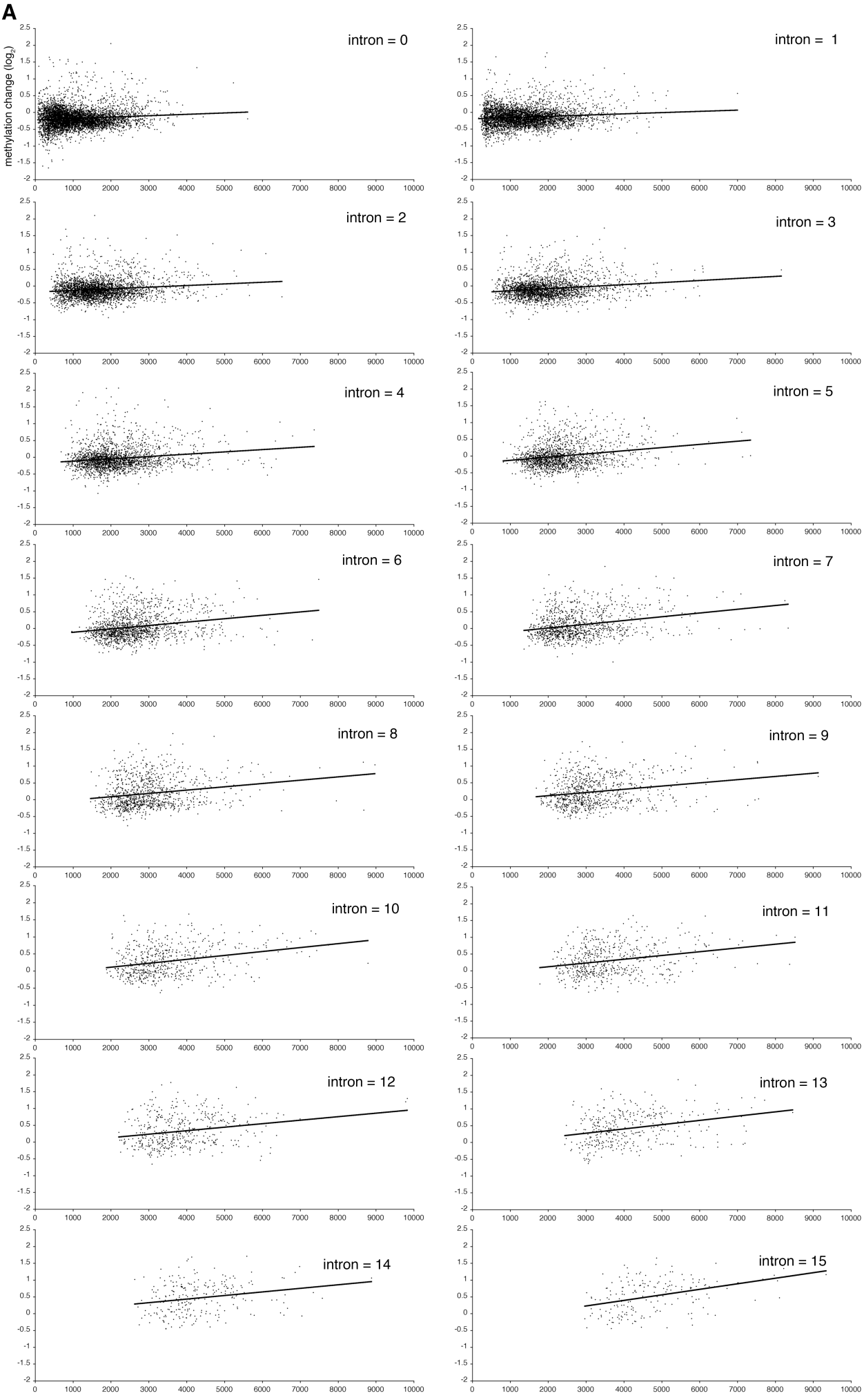
Supplementary Figure S14 The *ibm1* mutation affects 5' and 3' halves of a gene in a coordinated manner. Each gene was divided into halves and average change in the methylation signal was compared. For controls, the methylation change was also compared to upstream and downstream flanking regions of the same length.

A

	M	A	U
Class I	348	165	34
Class II	1173	642	217
Class III	2760	4235	11651

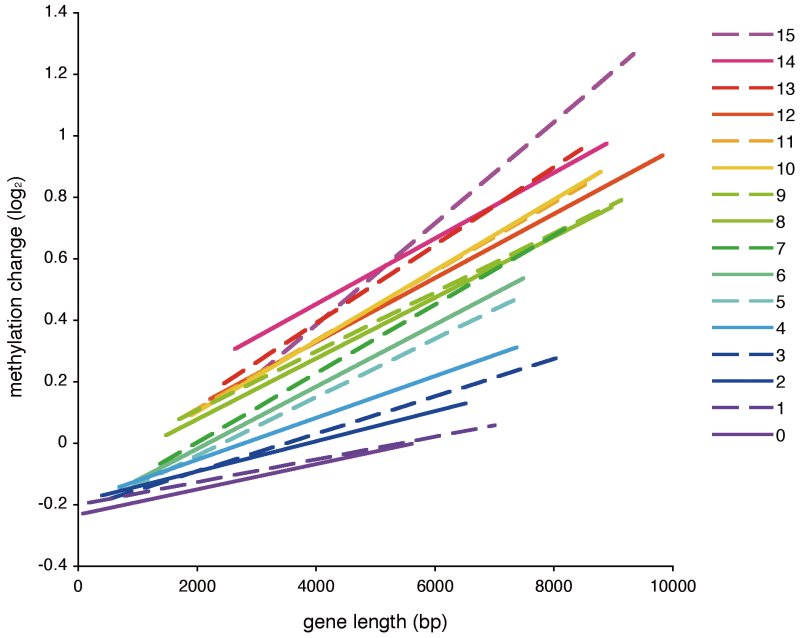


Supplementary Figure S15 (**A**) Relationship between DNA methylation in wild type and the response to the *ibm1* mutation. Methylated (M), ambiguous (A), and unmethylated (U) gene groups reflect methylation level in wild type, which are from Zilberman et al. (Nat Genet 39, 61-69). Gene Class I, II, and III are classified by response to the *ibm1* mutation (Supplementary Fig. S1). Gene number is indicated in each group. (**B**-**G**) Histograms showing the expression profiles of gene groups defined in a. Class I and II are combined. X-axis indicates $\log_{10}$ of expression in wild type leaves in our microarray experiment. Class III groups show bipartite pattern of expression, indicating that they include genes with high- and low-expression levels. In Class I and II, proportion of genes with low expression level was reduced. That was also seen in genes not methylated in wild type (A and U).

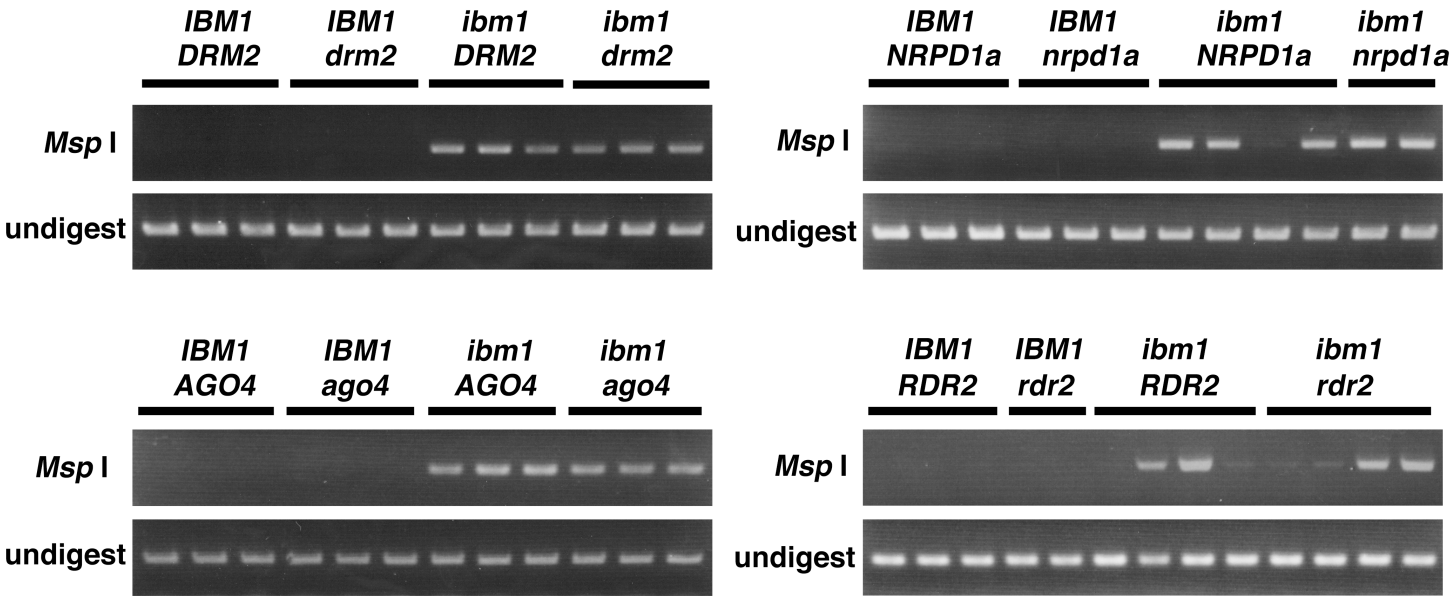


Supplementary Figure S16 (legend in the next page)

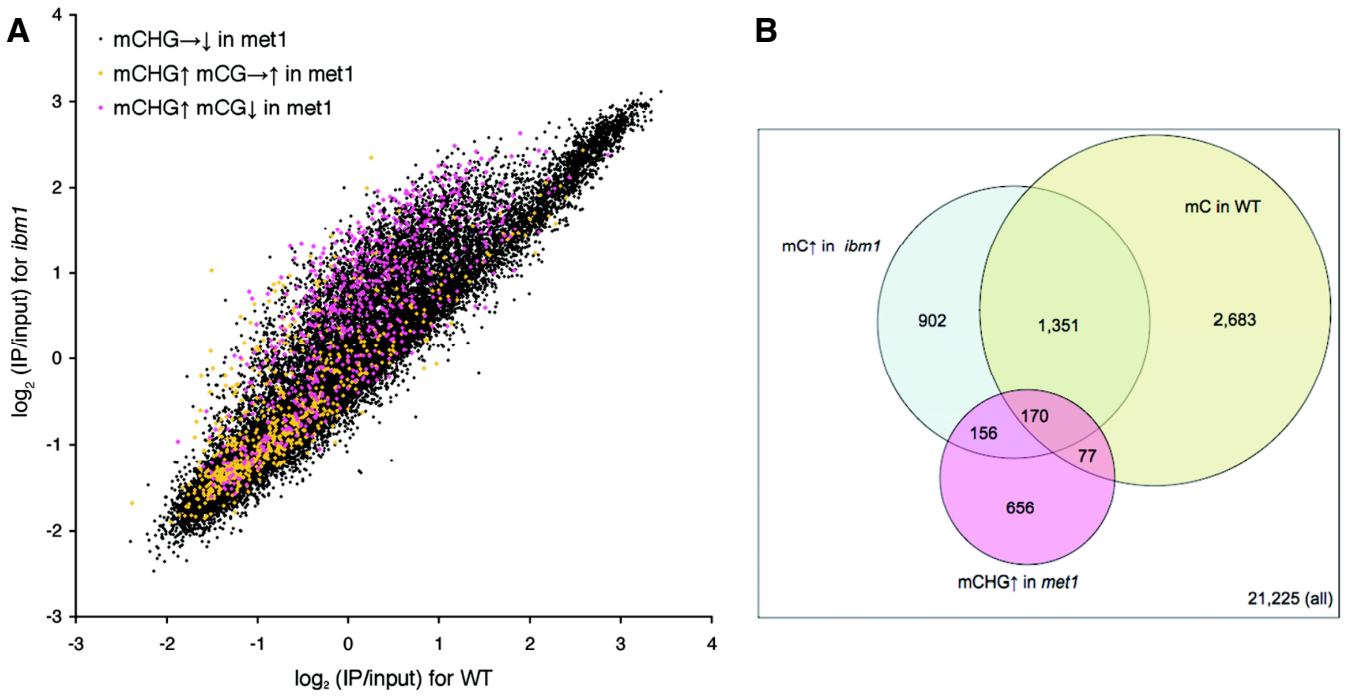
B



Supplementary Figure S16 Effect of intron number and gene length on the methylation change in *ibm1*. (**A**; previous page) Relationship between gene length and the methylation change in *ibm1*. Each graph shows the result of each gene group having the same intron number. The methylation change shows mean of $\log_2$ (ibm1/WT) signals in each gene. Lines were drawn by linear approximation. (**B**) Superimposed lines from (A). Even within a group of genes with the same intron number, longer gene showed stronger response to *ibm1*.



Supplementary Figure S17 Methylation status of *ERL2* gene (At5g07180) in the double mutants. Genomic DNA used in the experiment in Fig.6 was digested with or without the methylation-sensitive restriction enzyme *Msp* I (CCGG, sensitive to methylation in the first C) and amplified using primer pair 5'-CAGATCATGTAAGGCTTGTCTTATGC-3' + 5'-ACGTACAGATTGGCAAGACTGAAG-3'.
Both the *Msp* I-digested and undigested samples were digested with *Dra* I and *Pst* I before the PCR. Although the *ibm1*-induced hypermethylation was less penetrant than that in the *BNS* locus, the double mutants also showed hypermethylation in the *ERL2* gene.



Supplementary Figure S18 Relationship between the CHG hyper-methylation in *met1* and *ibm1*. (A) Methylation signal for each gene in wild type and *ibm1*. Genes showing increased CHG methylation and decreased CG methylation in *met1* tend to be hyper-methylated in *ibm1* (magenta dots). On the other hand, the response to *ibm1* was not so strong in genes showing increased CHG methylation without decrease in CG methylation in *met1* (yellow dots). (B) Venn diagram showing the relationship between genes methylated in wild type (yellow circle; Zilberman et al. 2006), genes with CHG hyper-methylation in *met1* (red circle; Lister et al. 2008), and genes hyper-methylated in *ibm1* (Blue circle; this work, Class I and Class II genes). Number of genes methylated in wild type and CHG hyper-methylation in *met1* (the overlap between yellow and red circles) was 247. Of the 247 genes, 170 showed hyper-methylation in *ibm1*.