Identification of DNA methylated regions by using methylated DNA immunoprecipitation sequencing in Brassica rapa

Takahashi, Satoshi ; Fukushima, Naoki ; Osabe, Kenji ; Itabashi, Etsuko ; Shimizu, Motoki ; Miyaji , Naomi ; Takasaki-Yasuda, Takeshi ; Suzuki…

(Citation)
Crop and Pasture Science,69(1):107-120

(Issue Date)
2018-01-04

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90005486

Title: Identification of DNA methylated regions using Methylated DNA immunoprecipitation sequencing in *Brassica rapa* L.

Satoshi Takahashi[A, H], Naoki Fukushima[B, H], Kenji Osabe[C, H], Etsuko Itabashi[D], Motoki Shimizu[E], Naomi Miyaji[B], Takeshi Takasaki-Yasuda[B], Yutaka Suzuki[F], Motoaki Seki[A, G], Ryo Fujimoto[B, I]

[A] RIKEN Center for Sustainable Resource Science, Yokohama, Kanagawa 230-0045, Japan

[B] Graduate School of Agricultural Science, Kobe University, Rokkodai, Nada-ku, Kobe 657-8501, Japan

[C] Plant Epigenetics Unit, Okinawa Institute of Science and Technology Graduate University, Onna-son, Okinawa 904-0495, Japan

[D] Institute of Vegetable and Floriculture Science, NARO, Kusawa, Ano, Tsu, Mie 514-2392, Japan

[E] Iwate Biotechnology Research Center, Narita, Kitakami, Iwate 024-0003, Japan

[F] Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba 277-0882, Japan

[G] Core Research for Evolutional Science and Technology, Japan Science and Technology, Kawaguchi, Saitama 332-0012, Japan.

[H] These authors equally contributed to this work

[I] Corresponding Author: Ryo Fujimoto

Graduate School of Agricultural Science, Kobe University, Rokkodai, Nada-ku, Kobe 657-8501, Japan

Phone: +81-78-803-5827

Fax: +81-78-803-5827

E-mail: leo@people.kobe-u.ac.jp

Running head: Identification of DNA methylated regions

Abstract

DNA methylation is an epigenetic gene regulatory mechanism that plays an essential role in gene expression, transposon silencing, genome imprinting, and plant development. We investigated the influence of DNA methylation on gene expression in *Brassica rapa*, to understand if there are epigenetic differences between inbred lines. Genome-wide DNA methylation was analyzed by Methylated DNA Immunoprecipitation sequencing (MeDIP-seq) of 14-day-old first and second leaves from two inbred lines of Chinese cabbage that are susceptible or resistant to Fusarium yellows. Model-based analysis for ChIP-seq (MACS) identified DNA methylation peaks in genic regions including 2 kb upstream, exon, intron, and 2 kb downstream regions. More than 65 % of genes showed similar patterns of DNA methylation in the genic regions in the two inbred lines. DNA methylation states of the two inbred lines were compared to their transcriptome. Genes having DNA methylation in the intron and the 200 bp upstream and downstream regions were associated with a lower expression level in both lines. A small number of genes showed a negative correlation between difference of DNA methylation levels and difference of transcriptional levels between the two inbred lines, suggesting that DNA methylation in these genes result in transcriptional suppression.

Additional keywords: DNA methylation, MeDIP-seq, Transposable elements, *Brassica rapa*, gene expression

Introduction

B. rapa shows morphological variation (morphotypes), and comprises

commercially important vegetable crops consumed worldwide including leafy vegetables such as Chinese cabbage (var. *pekinensis*), pak choi (var. *chinensis*), and komatsuna (var. *perviridis*), root vegetables including turnip (var. *rapa*), and oilseed (var. *oleifera*). Chinese cabbage forms a head with large pale-green colored leaves and wide white midribs and is an important vegetable in Asia. The reference genomes of *B. rapa* and its relative species of *B. oleracea* and *B. napus* are available (Wang *et al.* 2011; Chalhoub *et al.* 2014; Liu *et al.* 2014; Parkin *et al.* 2014), allowing detailed genetic and evolutionary studies.

Epigenetics can be associated with changes in the expression of the genome that do not involve changes in DNA sequence and epigenetic control is known to play an essential role in normal plant development (Fujimoto *et al.* 2012; Osabe *et al.* 2012; Matzke and Mosher 2014). DNA methylation is an epigenetic mark that adds a methyl group to the C-5 position of the cytosine ring (methylcytosine), and can be heritable and influence gene expression, transposon silencing, and genome imprinting. In plants, DNA methylation can occur in three different contexts of CG, CHG, and CHH (where H can be A, C, or T), and is regulated through different pathways (Fujimoto *et al.* 2012; Osabe *et al.* 2012; Matzke and Mosher 2014). DNA methylation can influence gene expression and affect plant phenotype, including agronomical traits. In *Arabidopsis thaliana*, DNA methylation that occurs in transcribed regions (gene-body methylation) was associated with high expression levels, whereas genes that are methylated in their promoter regions tended to show tissue-specific expression (Zhang *et al.* 2006). However, in rice, gene repression by DNA methylation in the transcriptional termination regions was stronger than the effect of DNA methylation in the promoter region (Li *et al.* 2012). The DNA methylation state at the whole genome level in *B. rapa* have been examined (Chen *et al.* 2015; Niederhuth *et al.* 2016), but there is no report showing the relationship between DNA methylation and expression levels within a plant or between different lines.

Increasing number of reports are revealing the association between DNA

methylation and agricultural traits. Flowering in Arabidopsis, fruit ripening in tomato, sex determination in melon, salt-tolerance in wheat, and drought tolerance in rice are some of the traits known to be epigenetically regulated or influenced (Kinoshita *et al.* 2006; Martin *et al.* 2009; Karan *et al.* 2012; Wang *et al.* 2014; Liu *et al.* 2015). In *B. napus* (rapeseed / canola), QTL analysis based on methylation sensitive amplified polymorphism (MSAP) revealed that 97 % of the methylation pattern of a particular parent line was stably inherited across at least 5 generations and some were linked to agronomical traits (epiQTL) (Long *et al.* 2011). From separate studies of *B. napus* that produced epigenetic recombinant inbred lines (epiRILs) that are epigenetically different but isogenic, artificial selection based on energy use efficiency was associated with particular epigenomic states that led to 5 % yield increase and drought tolerance (Hauben *et al.* 2009; Verkest *et al.* 2015). Treatment of *B. rapa* with 5-azaC, a cytidine analog that can inhibit DNA methylation, demonstrated male sterility, reduced seed size, and a late flowering phenotype, suggesting a strong relationship between DNA methylation and these traits (Amoah *et al.* 2012). Epigenetics in agriculture is becoming increasingly important but epigenetically regulated traits cannot be identified by conventional genomic studies, and cost-effective methods need to be developed to identify trait associated epialleles in various crop species.

Recent advances in sequencing technology allow us to investigate the epigenetic states at the genome-wide level, and methods such as WGBS (Whole genome bisulfite sequencing), MBD-seq (Methyl-CpG-binding domain sequencing), EpiRAD-seq (Epi-Restriction site associated DNA sequencing), and MeDIP-seq (Methylated DNA immunoprecipitation sequencing) have been developed for this purpose (Harris *et al.* 2010; Laird 2010; Schield *et al.* 2016). MeDIP-seq is a method to investigate the genome-wide methylation states by high-throughput sequencing enriched for methylated DNA fragments by immunoprecipitation using antibodies raised against methylcytosine. Methylation enriched fragments mapped against the genome will represent defined methylated regions (e.g. promoter, exons, introns etc.)

that can then be compared to other samples to identify the changes in DNA methylation. The methylation changes compared to phenotypic, transcriptomic, or proteomic data may help identify agronomically important epialleles that are regulated through DNA methylation.

Fusarium yellows (also known as Fusarium wilt) is caused by a soil-borne *Fusarium oxysporum* f. sp. *conglutinans* (*Foc*) or *F. oxysporum* f. sp. *rapae* in *Brassica* vegetables and is an economically important disease for Chinese cabbage (Enya *et al*. 2008). Leaf yellowing, wilting, defoliation, stunted growth, and death of seedlings are caused by infection of this pathogen, which invades the host roots and colonizes in their xylem tissues, especially in warm soil. A candidate resistance (*R*) gene against Fusarium yellows has been identified in *B. rapa* and *B. oleracea*, and they are orthologous and encodes a TIR-NBS-LRR protein (Lv *et al*. 2014; Shimizu *et al*. 2014, 2015). Resistant and susceptible lines in *B. rapa* have different immune responses against *Foc* inoculation. The resistant lines activates the genes involved in disease resistance such as 'Systemic acquired resistance', 'Regulation of defense response', and 'Response to salicylic acid stimulus' at 24 hours after inoculation (HAI) but not at 72 HAI or in susceptible lines at 24 and 72 HAI, suggesting that the defense response against *Foc* may be established by up-regulating these genes involved in resistance at 24 HAI in resistant lines (Miyaji *et al*. 2017).

In this study, to identify the DNA methylated region in *B. rapa*, we performed MeDIP-seq of two inbred lines of *B. rapa*, which show a difference in Fusarium yellows disease resistance caused by infection of *F. oxysporum* f. sp. *conglutinans* (Shimizu *et al.* 2014). DNA methylation states were similar between the two lines, but we identified regions that were specifically methylated in one of the lines. We examined the impact of DNA methylation on transcription by comparing the DNA methylation data to previous RNA sequencing (RNA-seq) data generated from samples using the same tissue and stage, but repeated independently (Shimizu *et al.* 2014). Genes having DNA methylation in the intron and the 200 bp upstream and downstream regions tended

to be repressed in both lines. A small number of genes showed a negative correlation between difference of DNA methylation levels and difference of transcription levels between the two inbred lines. The knowledge of DNA methylation state at the whole genome level will be useful for examining natural variation of DNA methylation states, change of DNA methylation states by abiotic or biotic stress, or understanding the contribution of DNA methylation to agronomically important traits using segregation of loci.

Materials and methods

Plant materials, DNA extraction, and RNA sequencing

Two Chinese cabbage inbred lines developed in a previous study, RJKB-T23 and RJKB-T24, were used as plant materials (Kawamura *et al.* 2016). Seven generations of selfing and selection based on traits concerned with the breeding objective has been performed in both inbred lines. Plants were grown in plastic dishes containing Murashige and Skoog (MS) agar medium supplemented with 1.0 % sucrose (pH5.7) in growth chambers under a 16-h/8-h light/dark cycle at 22 °C. Fourteen-day-old first and second leaves harvested from RJKB-T23 and RJKB-T24 were used for genomic DNA extraction. Total genomic DNA for MeDIP-seq or chop-PCR was isolated by the Cetyl trimethyl ammonium bromide method (Murray and Thompson 1980).

RNA-sequencing (RNA-seq) using 14-day-old first and second leaves has been performed previously without replication for RJKB-T23 and RJKB-T24 (Shimizu *et al*. 2014), and replication was conducted independently under the same biological (tissues, stages, and growth condition) and technical (50 nt read length with single end on an Illumina HiSeq^{TM} 2000) conditions. The two replicates showed high correlation, r=0.98 (RJKB-T23) and r=0.99 (RJKB-T24).

Methylated DNA immunoprecipitation (MeDIP)

MeDIP was performed as described previously using genomic DNAs of 14-day-old first and second leaves (Kawanabe *et al.* 2012). The genomic DNA was fragmented by sonication, to sizes ranging from 150 bp to 700 bp (peak size is about 300 bp). Anti-methylcytosine antibody (Diagenode, NJ, USA) was used to obtain purified immunoprecipitated DNAs. Enrichment of methylated DNA fragments in the immunoprecipitated DNAs was confirmed by qPCR using the regions that are known to be methylated (positive control), *BrTto1*, *BrSTF7a*, and *BrSTF12b* (Fujimoto *et al.* 2008a; Sasaki *et al.* 2011), and non-methylated (negative control), Bra001846 and Bra023446 (Table S1).

For qPCR, MeDIP-DNA was amplified using FastStart Essential DNA Green Master (Roche) using a LightCycler Nano (Roche). PCR conditions were 95°C for 10 min followed by 40 cycles of 95 °C for 10 s, 60 °C for 10 s, and 72 °C for 15 s, and Melting program (60 °C to 95 °C at 0.1 °C/s). After amplification cycles, each reaction was subjected to melt temperature analysis to confirm single amplified products. Data presented are the average and standard error (SE) from three biological and experimental replications. Ratio of amplification of positive control genes to negative control genes were compared between MeDIP-DNA and input-DNA as templates for confirmation of enrichment of methylated DNA by MeDIP.

Methylated DNA immunoprecipitation sequencing (MeDIP-seq)

MeDIP-seq was performed in two biological replicates for 36-bp single-end and 50-bp paired-end sequencing. Each replicates was sown on different days but grown under the same conditions and harvested at the same developmental stage. The samples developmental stages and growth conditions used for MeDIP-seq were the same as those of RNA-seq. We commissioned the second sequencing (50-bp paired-end) to Beijing Genomics Institute (BGI).

Samples of immunoprecipitated DNAs and Input-DNA were sequenced by Hiseq2000 (36-bp single-end or 50-bp paired-end) after PCR amplification and size

selection (200-300 bp). The reads of MeDIP-seq were purged from low quality reads or adapter sequences using cutadapt version 1.7.1 and Trim Galore! version 0.3.7. Then the reads were mapped to the *B. rapa* reference genome v.1.5 using Bowtie2 version 2.2.3. We performed peak calling on alignment results using Model-based analysis for ChIP-seq (MACS) 2 2.1.0 and identified the regions having DNA methylation as peaks. The MACS callpeak was used with the following options (effective genome size: 2.30e+08, band width: 200, model fold: 10-30, tag size: 36). The cutoff of p-value, 1.00e-05, was used to call significant peaks.

To estimate the difference of methylated genic regions as peak basis between RJKB-T23 and RJKB-T24, the total numbers of methylated genic regions were counted when their total length was over 200 bp and counterpart was 0 bp.

To statistically estimate the difference of methylated genic regions between RJKB-T23 and RJKB-T24 using Reads Per Million (RPM) score, a target region that contains a gene, 200 bp upstream and 200 bp downstream was used. The target region was divided equally into 30 divisions as windows. The number of reads mapped to a window was counted and normalized to RPM score for MeDIP-seq and Input-DNA-seq. To normalize the RPM of a window, we subtracted Input-DNA RPM value from MeDIP RPM value for each window. T-statistic of a region was calculated using the difference of RJKB-T23 window's RPM and RJKB-T24 window's RPM. Statistical significance of differences between RJKB-T23 and RJKB-T24 was determined by one-sample t-test. The regions that showed significant differences were selected with q-value < 0.05 and average of window's RPM > 0.3 on either sample.

Chop-PCR

Chop-PCR experiment was performed as described by Kawanabe *et al.* 2016. Fifty ng of genomic DNA was digested with *Hpa* II in 20μl reaction mix at 37°C for five hours. After restriction digestion, 1μl of digested DNA was used as template for PCR in 10μl reaction mix. The PCR conditions were 94 °C for 2 min followed by 35

cycles of 94 °C for 30 s, 55 °C for 30 s, and 68 °C for 30 s. Primers used for chop-PCR are listed in Table S1.

Gene Ontology analysis

Analysis for enrichment of gene functional ontology terms was completed using the gene ontology (GO) tool agriGO (Du *et al.* 2010) following the methods described by Shimizu *et al.* 2014. Statistical tests for enrichment of functional terms used the hypergeometric test and false discovery rate (FDR) correction for multiple testing to a level of 1 % FDR.

Results

Methylated DNA Immunoprecipitation sequencing

To identify the DNA methylated regions in *B. rapa*, we performed MeDIP-seq analysis using 14-day-old first and second leaves of Chinese cabbage inbred lines, RJKB-T23 and RJKB-T24. We used HiSeq2000 (36 bp single-end) for sequencing, and 87,710,099 and 53,946,789 clean reads were obtained from Input-DNA-seq of RJKB-T23 and RJKB-T24, respectively, and 42,128,140 (48.0 %) and 27,327,630 reads (50.7 %) from RJKB-T23 and RJKB-T24, respectively, were uniquely mapped to the *B. rapa* reference genome (Table 1). From MeDIP-seq with 36 bp single-end, 23,250,396 and 25,398,065 clean reads were obtained in RJKB-T23 and RJKB-T24, respectively, and 6,277,802 (27.0 %) and 6,645,462 reads (26.2 %) in RJKB-T23 and RJKB-T24, respectively, were uniquely mapped to the *B. rapa* reference genome (Table 1).

We also sequenced MeDIP-DNA using HiSeq2000 (50 bp paired-end) as a replicate, and 13,355,866 and 14,081,034 clean reads were obtained of which 9,610,672 (72.0 %) and 10,096,202 (71.7 %) reads in RJKB-T23 and RJKB-T24, respectively, were uniquely mapped to the *B. rapa* reference genome (Table 1). 50 bp paired-end reads were mapped more successfully than the 36 bp single-end reads (Table 1). However, chromosomal distribution of mapped reads analyzed by the sliding window of

100 kb was similar between the replicates in both RJKB-T23 and RJKB-T24 (Fig. 1, Fig. S1, S2), and log 10 score of reads per kilobase of exon per million mapped reads (RPKM) in each window correlated significantly between the replicates in both RJKB-T23 and RJKB-T24 (Fig. 1). We combined the data from the replicates.

It is well known that DNA methylation is observed in repetitive sequences such as transposable elements (TEs), so we examined the mapped reads on the interspersed repeats regions (IRRs) such as TEs detected by RepeatMasker. In both lines, the percentages of mapped reads on the IRRs using MeDIP-seq data were higher than those using Input-DNA-seq data (Fig. S3, Table 1).

We classified the mapped reads on the genic region into four categories, 2 kb upstream, exon, intron, and 2 kb downstream, using Input-DNA-seq and MeDIP-seq data of RJKB-T23 and RJKB-T24. The proportions of mapped reads in these four categories using MeDIP-seq data were lower than that in Input-DNA-seq in both lines (Table 2). Proportion of mapped reads of MeDIP-seq in the 2 kb upstream and exon regions were higher and lower than those of Input-DNA-seq, respectively, in both lines (Fig. S4, Table 2).

Detection of the peaks of methylated regions

Model-based analysis for ChIP-seq (MACS) was used for scanning the DNA methylation peaks, and 45,558 and 49,142 DNA methylation peaks were identified in RJKB-T23 and RJKB-T24, respectively (Table 3). 39,797 (87.4 %) and 42,490 peaks (86.5 %) were found in the IRRs in RJKB-T23 and RJKB-T24, respectively (Table 3). We counted the number of genes having DNA methylation peaks in RJKB-T23, and 8,709, 1,950, 2,144, and 6,595 genes had more than one DNA methylation peak within 2 kb upstream, exon, intron, and 2 kb downstream regions, respectively (Table 3). In RJKB-T24, 8,474, 2,268, 2,313, and 6,671 genes had more than one DNA methylation peak within 2 kb upstream, exon, intron, and 2 kb downstream regions, respectively (Table 3). More than 66 % of genes having DNA methylation peaks overlapped between

the two lines (Fig. S5).

More than 79 % of peaks in the 2 kb upstream, intron, or 2 kb downstream regions overlapped with the peaks in the IRRs, while 60 % of peaks in the exon regions overlapped with the peaks in the IRRs in both RJKB-T23 and RJKB-T24 (Table 3). Most of the top 20 longest DNA methylation peaks were observed in the IRRs of intergenic regions of RJKB-T23 and RJKB-T24 (Fig. S6a, Table S2). We identified the top 20 longest DNA methylation peaks harboring genic regions (2 kb upstream, exon, intron, and 2 kb downstream), and most of the DNA methylation peaks overlapped with IRRs such as retrotransposons (*copia*- or *gypsy*-type) and DNA type transposons (*En-Spm* or *MuDR*) (Fig. S6b, Table S3).

Genes having DNA methylation and their expression levels

The level of gene expression of the transcriptomes of 14-day-old first and second leaves of RJKB-T23 and RJKB-T24 (Shimizu *et al.* 2014) were categorized into seven groups using log2 score of fragments per kilobase of transcript per million mapped reads (FPKM) in RJKB-T24, e.g., Group-6 (highest), log2 score of FPKM (x) is greater than 9.0; Group-5, 6.0<=x<9.0; Group-4, 3.0<=x<6.0; Group-3, 0.0<=x<3.0; Group-2, -3.0<=x<0.0; Group-1, x<-3.0; Group-0, no read (lowest) (Kawanabe *et al.* 2016), and in this study we categorized the gene expression levels in RJKB-T23 using the same criteria (Fig. S7). We classified genes having DNA methylation peaks in 2 kb upstream, exon, intron, and 2 kb downstream regions into these 7 groups of expression levels in RJKB-T23 and RJKB-T24. Of genes having DNA methylation in 2 kb upstream or downstream region, the distribution from group-0 to -6 was similar to that in total genes in both lines (Fig. S7). By contrast, the genes having DNA methylation in exon or intron region were over Group-0 in both lines (Fig. S7). The average of expression levels of genes (log2 score of FPKM, FPKM > 0.01) having a DNA methylation peak in the 2 kb upstream or downstream region was similar to that in total genes in both lines, while the average expression level of genes having a DNA

methylation peak in the exon and intron regions was lower than that of total genes (Fig. 2).

We calculated the RPKM using mapped reads of MeDIP-seq in six regions (2 kb upstream, 200 bp upstream, exon, intron, 200 bp downstream, 2 kb downstream). The correlation coefficient between methylation levels (log2 score of RPKM, RPKM > 0.01) in each region and the expression levels (log2 score of FPKM, FPKM > 0.01) was examined. There was a negative correlation between the methylation levels and expression levels in the 200 bp upstream, intron, 200 bp downstream, or 2 kb downstream regions (Table 4), indicating that DNA methylation in these regions results in the repression of expression.

We selected genes having DNA methylation peaks in both exon and intron regions or in all four regions (2 kb upstream, exon, intron, and 2 kb downstream) in RJKB-T23 and RJKB-T24. 1,212 and 1,403 genes had DNA methylation peaks in both exon and intron regions in RJKB-T23 and RJKB-T24, respectively, and we performed a Gene Ontology (GO) analysis of these genes. Twenty-four and 33 GO categories were significantly overrepresented in RJKB-T23 and RJKB-T24, respectively, and the GO categories of ‘Catalytic activity’, ‘Post-embryonic development’, ‘Hydrolase activity’, ‘CUL4 RING ubiquitin ligase complex’, and ‘Nucleotide binding’ were significantly overrepresented in both RJKB-T23 and RJKB-T24 (Table S4). The 394 and 481 genes that had DNA methylation peaks in all four regions of RJKB-T23 and RJKB-T24, respectively, were heavily methylated. GO analysis of these heavily methylated genes was performed, and none of the GO category was significantly overrepresented in both lines.

Validation of DNA methylation by chop-PCR

We confirmed the results of MeDIP-seq in the regions by chop-PCR in both RJKB-T23 and RJKB-T24. We assessed ten regions (eight regions having DNA methylation and two regions without) by chop-PCR using the DNA methylation

sensitive restriction enzyme *Hpa* II. All eight regions having DNA methylation showed PCR amplification, while two regions not having DNA methylation showed no amplification (Fig. 3).

Comparison of the DNA methylated regions between two inbred lines

The DNA methylation states between RJKB-T23 and RJKB-T24 were compared by two methods. First, we compared the DNA methylation states in the genic regions (2 kb upstream, exon, intron, and 2 kb downstream) using the data of DNA methylation peaks. 1,756 DNA methylation peaks were observed only in RJKB-T23, but not in RJKB-T24 (termed T23-SMG; T23 specifically methylated genes), and 1,870 DNA methylation peaks were observed only in RJKB-T24, but not in RJKB-T23 (T24-SMG) (Fig. 4, 5, Fig. S8, Table S5). The regions having differential DNA methylated peaks were observed in the 2 kb upstream and downstream regions rather than exon or intron regions (Fig. 5, Table S5). We examined whether these differential DNA methylation peaks affect the gene expression level using previous RNA-seq data (Shimizu *et al*. 2014). In T23-SMG, 43 of 1,621 genes (2.7 %) showed differential expression, and 19 and 24 genes showed a higher and lower expression level in RJKB-T23 than in RJKB-T24, respectively (Table 5). In T24-SMG, 27 of 1,705 genes (1.6 %) showed differential expression, and 10 and 17 showed a higher and lower expression level in RJKB-T23 than in RJKB-T24, respectively (Table 5). 478 of 1,621 T23-SMG (30.0 %) and 623 of 1,705 T24-SMG (36.5 %) were not expressed in either RJKB-T23 and RJKB-T24.

The differentially methylated regions between RJKB-T23 and RJKB-T24 were also identified using a comparison of reads per million mapped reads (RPM) scores (see materials and methods). 447 genes showed higher DNA methylation levels in RJKB-T23 than RJKB-T24 (termed T23-HMG) (Table S6, Figure S8), and one and three genes showed a higher and lower expression level in RJKB-T23 than in RJKB-T24, respectively (Table 6). 896 genes showed higher DNA methylation levels in

RJKB-T24 than RJKB-T23 (T24-HMG) (Table S6, Figure S8), and seven and two genes showed a higher and lower expression level in RJKB-T23 than in RJKB-T24, respectively (Table 6). More than 70 % of differentially methylated genes were not expressed in both RJKB-T23 and RJKB-T24.

These two analyses revealed that a small number of genes showed negative correlation between a difference of DNA methylation levels and expression levels in RJKB-23 and RJKB-T24, and many genes having differentially DNA methylated states between RJKB-T23 and RJKB-T24 were not expressed.

Using T23- and T24-HMG, we performed a GO analysis. 27 and 47 GO categories were significantly overrepresented in T23- and T24-HMG, respectively, and 24 GO categories such as ‘Metabolic process’, ‘Catalytic activity’, ‘Oxidation reduction’, and ‘Nucleotide binding’ overlapped in both T23- and T24-HMG (Table S7). GO categories of ‘Ion binding’, ‘Integral to membrane’, and ‘Structural constituent of cell wall’ were specifically overrepresented in T24-HMG (Table S7).

Identification of genes that have DNA methylation under normal condition and their gene expression changes by Foc inoculation

We have identified genes whose expression changed by *Foc* inoculation both in Fusarium yellow resistant line, RJKB-T23, and susceptible line, RJKB-T24 (Miyaji *et al*. 2017). We examined whether these genes had DNA methylation peaks based on the data of MeDIP-seq produced in this study. Of 260 differentially expressed genes between *Foc*- and mock-inoculated samples at 24 hours after inoculation (HAI) in RJKB-T23, 98 (37.7 %) genes had DNA methylation peaks in genic regions (2 kb upstream, exon, intron, or 2 kb-downstream) (Table S8). Of 253 differentially expressed genes at 24 HAI in RJKB-T24, 87 (34.4 %) genes had DNA methylation peaks in genic regions, and 36 genes were common between both lines (Table S8). In the resistant line, some genes involved in defense response such as *ACO1* (*ACC OXIDASE 1*), *BGLU18* (*BETA GLUCOSIDASE 18*), *Chitinase*, *ELI3* (*ELICITOR-ACTIVATED GENE 3*),

GSTF3 (*GLUTATHIONE S-TRANSFERASE F3*), *HIR2* (*HYPERSENSITIVE INDUCED REACTION 2*), *JAZ1 (JASMONATE-ZIM-DOMAIN PROTEIN 1)*, *NDR1* (*NON RACE-SPECIFIC DISEASE RESISTANCE 1*), *RBOHD* (*RESPIRATORY BURST OXIDASE HOMOLOGUE D*), *PR-3* (*PATHOGENESIS-RELATED 3*), and *WRKY51*, were up-regulated by *Foc* inoculation at 24 HAI and had corresponding DNA methylation peaks.

Discussion

We performed MeDIP-seq using two inbred lines of Chinese cabbage to identify the methylated regions of the DNA. We compared the percentage of multiple and unique mapped reads of 36 bp single-end sequencing runs between Input-DNA-seq and MeDIP-seq, and the percentages in MeDIP-seq were lower than those in Input-DNA-seq in both lines, suggesting that unmapped reads of MeDIP-seq had been omitted. In *B. rapa*, genome sequences of most euchromatic regions were determined, which is approximately half of the total genome size, but the sequences of the heterochromatic regions including the centromeres or pericentromeres were not determined (Wang *et al.* 2011). This indicates that the unmapped reads of MeDIP-seq derived from the heterochromatic regions. DNA methylation is enriched in heterochromatic regions of genomes, which consist of repetitive sequences and TEs in *A. thaliana* (Cokus *et al.* 2008; Lister *et al.* 2008; Zhang *et al.* 2008). The presence of DNA methylation in some TEs or repetitive sequences and the association between higher levels of DNA methylation and enrichment of repetitive sequences has been reported in *B. rapa* (Fujimoto *et al.* 2008a; Sasaki *et al.* 2011; Chen *et al.* 2015). In this study, more reads were mapped on the IRRs in MeDIP-seq than in Input-DNA-seq, indicating that IRRs in euchromatic regions were highly methylated in *B. rapa*.

The DNA methylated regions in genic regions including not only exon and intron regions, but also 2 kb upstream and 2 kb downstream regions were examined. Among four regions, more DNA methylation peaks were detected in the 2 kb upstream

and downstream regions than in exon regions, and genes having DNA methylation peaks tended to have IRRs, suggesting that detection of DNA methylation peaks in the genic regions was due to the DNA methylation in IRRs. This suggests the higher percentage of DNA methylation peaks in the 2 kb upstream and 2 kb downstream regions was due to the higher frequency of IRRs in the 2 kb upstream and 2 kb downstream regions than in the exon regions. We performed GO analysis using heavily methylated genes having DNA methylation peaks throughout the genic regions but no category was overrepresented, indicating that DNA methylated region and gene function are independent.

DNA methylations in TEs around or within the genic regions can affect gene expression levels (Liu *et al.* 2004; Saze *et al.* 2008; Martin *et al.* 2009; Fujimoto *et al.* 2012). The average of gene expression levels in the genes having DNA methylation peaks in the 2 kb upstream and 2 kb downstream regions was similar to that of total genes, but the average of gene expression levels in the genes having DNA methylation peaks in exon and intron regions were lower than that of total genes. In addition, more genes having DNA methylation peaks in exon and intron regions were not expressed. Gene expression levels and DNA methylation levels using RPKM scores in the six regions showed a negative correlation in the 200 bp upstream and intron regions, and strong and weak negative correlations in the 200 bp and 2 kb downstream regions, respectively. The genes having only CG methylation in the exon region, termed gene body methylation, show moderate gene expression levels in many plant species. About 14 % of genes have gene body methylation in *A. thaliana*, however, only 0.5 % of genes had gene body methylation in *B. rapa* (Niederhuth *et al.* 2016). The proportion of genes having DNA methylation peaks in exon regions was smaller and genes having DNA methylation peaks in exon regions showed low expression level. This may be due to a small fraction of gene body methylation with moderate gene expression levels and/or preferential detection of densely methylated regions by MeDIP-seq. Thus, the discrepancy in the exon regions between the two analyses (peak and RPKM based) is

considered as a mixture of gene body methylation with moderate gene expression level and methylation inducing silencing when calculating the RPKM. From these two analyses, we consider that DNA methylation in intron regions and 200 bp upstream and downstream regions results in silencing of gene expression.

We compared the DNA methylation states between two inbred lines by two analyses. We identified regions having differential DNA methylation peaks between the two inbred lines and examined the effect on gene expression levels. Most genes having different DNA methylation peaks between the two lines showed similar gene expression levels and about 30% of genes were not expressed. In T23-SMG and T24-SMG, which is specifically methylated in the genic region of one line, some genes showed differential expression between the two lines, but higher DNA methylation states did not cause lower expression levels. We also detected differentially methylated regions between the two inbred lines using RPM scores. Among differentially methylated genes, there were several genes showing differences of gene expression levels, while more than 70 % of differentially methylated genes were not expressed. There is a weak negative correlation between differences in DNA methylation and differences in gene expression between accessions of rice or *A. thaliana* (Zhang *et al.* 2008; He *et al.* 2010), while other studies have reported no relationship between accessions of rice or *A. thaliana* (Vaughn *et al.* 2007; Li *et al.* 2012). There is some evidence that DNA methylation in specific regions, especially in the promoter regions, represses gene expression (Saze and Kakutani 2007; Fujimoto *et al.* 2008b, 2011; Tarutani *et al.* 2010). In this study, we found a few genes showing a negative correlation between differences in DNA methylation and difference in gene expression between the two inbred lines, suggesting that these genes might be regulated by DNA methylation.

In this study we identified differentially methylated regions between two inbred lines, which have different disease resistance against Fusarium yellows (Shimizu *et al.* 2014). In *A. thaliana*, it has been shown that DNA methylation plays an important role in disease resistance, and several hypomethylated mutants enhanced disease

resistance (Zhu *et al.* 2016). Mutants in the genes involved in DNA demethylase have shown increased susceptibility to the fungal pathogen *F. oxysporum* (Le *et al.* 2014). In addition, DNA methylation states were globally changed in response to biotic stress (Dowen *et al.* 2012, Zhu *et al.* 2016). We identified differentially methylated regions between two lines in normal growth condition, and we did not find any changes in the GO categories related to biotic stress. However, we identified the genes that have DNA methylation and changed their expression levels in response to *Foc* inoculation in Fusarium yellows resistant or susceptible line, and some of them are related to defense response and up-regulated in the resistant line. Although we did not examine the DNA methylation state after *Foc* inoculation in this study, *Foc* inoculation may alter DNA methylation state and expression of defense responsive genes. Thus, our MeDIP-seq analysis might be useful to examine the change of DNA methylation states that occur between the resistant and susceptible lines to identify the loci involved during or after *Foc* inoculation.

There are many approaches for examining DNA methylation state at the whole genome level that have their own strengths and weaknesses, and the choice of method depends on the number of samples, quality and quantity of DNA, or desired coverage and resolution (Laird 2010). The most comprehensive method to detect methylated regions at the whole genome is WGBS, which provides methylation data at the single base resolution, but this method requires more sequence reads compared with MeDIP-seq. Thus, MeDIP-seq provides information about methylated genomic regions at a fraction of the cost of WGBS. In this study, we identified the methylated genomic regions and differentially methylated regions between the two lines, suggesting that MeDIP-seq is sufficient for producing meaningful results. Further study will be required to confirm that gene expression is regulated by DNA methylation by using a DNA methyltransferase inhibitor or hypomethylated transgenic plants or mutants (Fujimoto *et al.* 2008a; Amoah *et al.* 2012), and assessing the impact on phenotypic variation such as biotic stress.

Acknowledgements

We thank Dr. Elizabeth S. Dennis for her helpful comments and editing manuscript. This work was supported by MEXT KAKENHI (No. 221S0002), by an Open Partnership Joint Projects of JSPS Bilateral Joint Research Projects (14544567), and by Grant-in-Aid for Young Scientists (B) (2478002), Scientific Research on Innovative Areas (24113509), and Grant-in-Aid Young Scientists (A) (15H05614) (JSPS) to R. Fujimoto, and grants from RIKEN and Japan Science and Technology Agency (JST) [Core Research for Evolutionary Science and Technology (JPMJCR13B4) (CREST)] to M. Seki.

References

Amoah S, Kurup S, Rodriguez Lopez CM, Welham SJ, Powers SJ, Hopkins CJ, Wilkinson MJ, King GJ (2012) A hypomethylated population of *Brassica rapa* for forward and reverse epi-genetics. *BMC Plant Biology* **12,** 193. doi: 10.1186/1471-2229-12-193.

Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X, Chiquet J, Belcram H, Tong C, Samans B, Corréa M, Da Silva C, Just J, Falentin C, Koh CS, Le Clainche I, Bernard M, Bento P, Noel B, Labadie K, Alberti A, Charles M, Arnaud D, Guo H, Daviaud C, Alamery S, Jabbari K, Zhao M, Edger PP, Chelaifa H, Tack D, Lassalle G, Mestiri I, Schnel N, Le Paslier MC, Fan G, Renault V, Bayer PE, Golicz AA, Manoli S, Lee TH, Thi VHD, Chalabi S, Hu Q, Fan C, Tollenaere R, Lu Y, Battail C, Shen J, Sidebottom CHD, Wang X, Canaguier A, Chauveau A, Bérard A, Deniot G, Guan M, Liu Z, Sun F, Lim YP, Lyons E, Town CD, Bancroft I, Wang X, Meng J, Ma J, Pires JC, King GJ, Brunel D, Delourme R, Renard M, Aury JM, Adams KL, Batley J, Snowdon RJ, Tost J, Edwards D, Zhou Y, Hua W, Sharpe AG, Paterson AH, Guan C, Wincker P (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* **345,** 950-953. doi:

10.1126/science.1253435.

Chen X, Ge X, Wang J, Tan C, King GJ, Liu K (2015) Genome-wide DNA methylation profiling by modified reduced representation bisulfite sequencing in *Brassica rapa* suggests that epigenetic modifications play a key role in polyploid genome evolution. *Frontiers in Plant Science* **6,** 836. doi: 10.3389/fpls.2015.00836.

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452,** 215-219. doi: 10.1038/nature06745.

Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR (2012) Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences of the United States of America* **109,** E2183-E2191. doi: 10.1073/pnas.1209329109.

Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: A GO analysis toolkit for the agricultural community. *Nucleic Acids Research* **38,** W64-W70. doi: 10.1093/nar/gkq310.

Enya J, Togawa M, Takeuchi T, Yoshida S, Tsushima S, Arie T, Sakai T (2008) Biological and phylogenetic characterization of *Fusarium oxysporum* complex, which causes yellows on *Brassica* spp., and proposal of *F. oxysporum* f. sp. *rapae*, a novel forma specialis pathogenic on *B. rapa* in Japan. *Phytopathology* **98,** 475–483. doi: 10.1094/PHYTO-98-4-0475.

Fujimoto R, Kinoshita Y, Kawabe A, Kinoshita T, Takashima K, Nordborg M, Nasrallah ME, Shimizu KK, Kudoh H, Kakutani T (2008b) Evolution and control of imprinted *FWA* genes in the genus *Arabidopsis*. *PLoS Genetics* **4,** e1000048. doi: 10.1371/journal.pgen.1000048.

Fujimoto R, Sasaki T, Inoue H, Nishio T (2008a) Hypomethylation and transcriptional reactivation of retrotransposon-like sequences in *ddm1* transgenic plants of *Brassica rapa*. *Plant Molecular Biology* **66,** 463-473. doi:

10.1007/s11103-007-9285-1.

Fujimoto R, Sasaki T, Ishikawa R, Osabe K, Kawanabe T, Dennis ES (2012) Molecular mechanisms of epigenetic variation in plants. *International Journal of Molecular Science* **13,** 9900-9922. doi: 10.3390/ijms13089900.

Fujimoto R, Sasaki T, Kudoh H, Taylor JM, Kakutani T, Dennis ES (2011) Epigenetic variation in the *FWA* gene within the genus Arabidopsis. *The Plant Journal* **66,** 831-843. doi: 10.1111/j.1365-313X.2011.04549.x.

Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y, Olshen A, Ballinger T, Zhou X, Forsberg KJ, Gu J, Echipare L, O'Geen H, Lister R, Pelizzola M, Xi Y, Epstein CB, Bernstein BE, Hawkins RD, Ren B, Chung WY, Gu H, Bock C, Gnirke A, Zhang MQ, Haussler D, Ecker JR, Li W, Farnham PJ, Waterland RA, Meissner A, Marra MA, Hirst M, Milosavljevic A, Costello JF (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnology* **28,** 1097-105. doi: 10.1038/nbt.1682.

Hauben M, Haesendonckx B, Standaert E, Van Der Kelen K, Azmi A, Akpo H, Van Breusegem F, Guisez Y, Bots M, Lambert B, Laga B, De Block M (2009) Energy use efficiency is characterized by an epigenetic component that can be directed through artificial selection to increase yield. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 20109-20114. doi: 10.1073/pnas.0908755106.

He G, Zhu X, Elling AA, Chen L, Wang X, Guo L, Liang M, He H, Zhang H, Chen F, Qi Y, Chen R, Deng XW (2010) Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. *The Plant Cell* **22,** 17-33. doi: 10.1105/tpc.109.072041.

Karan R, DeLeon T, Biradar H, Subudhi PK (2012) Salt stress induced variation in DNA methylation pattern and its influence on gene expression in contrasting rice genotypes. *PLoS One* **7,** e40203. doi: 10.1371/journal.pone.0040203.

Kawamura K, Kawanabe T, Shimizu M, Nagano AJ, Saeki N, Okazaki K, Kaji M, Dennis ES, Osabe K, and Fujimoto, R (2016) Genetic distance of inbred lines of Chinese cabbage and its relationship to heterosis. *Plant Gene* **5,** 1-7. doi:10.1016/j.plgene.2015.10.003

Kawanabe T, Fujimoto R, Sasaki T, Taylor JM, Dennis ES (2012) A comparison of transcriptome and epigenetic status between closely related species in the genus *Arabidopsis*. *Gene* **506,** 301-309. doi: 10.1016/j.gene.2012.07.003.

Kawanabe T, Osabe K, Itabashi E, Okazaki K, Dennis ES, Fujimoto R (2016) Development of primer sets that can verify the enrichment of histone modifications, and their application to examining vernalization-mediated chromatin changes in *Brassica rapa* L. *Genes & Genetic Systems* **9,** 1-10. doi: 10.1266/ggs.15-00058.

Kinoshita Y, Saze H, Kinoshita T, Miura A, Soppe WJJ, Koornneef M, Kakutani T (2006) Control of *FWA* gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. *The Plant Journal* **49,** 38-45. doi: 10.1111/j.1365-313X.2006.02936.x

Laird PW (2010) Principles and challenges of genome-wide DNA methylation analysis. *Nature Reviews Genetics* **11,** 191-203. doi: 10.1038/nrg2732.

Le TN, Schumann U, Smith NA, Tiwari S, Au PC, Zhu QH, Taylor JM, Kazan K, Llewellyn DJ, Zhang R, Dennis ES, Wang MB (2014) DNA demethylases target promoter transposable elements to positively regulate stress responsive genes in *Arabidopsis*. *Genome Biology* **15,** 458. doi: 10.1186/s13059-014-0458-3.

Li X, Zhu J, Hu F, Ge S, Ye M, Xiang H, Zhang G, Zheng X, Zhang H, Zhang S, Li Q, Luo R, Yu C, Yu J, Sun J, Zou X, Cao X, Xie X, Wang J, Wang W (2012) Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics* **13,** 300. doi: 10.1186/1471-2164-13-300

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in

Arabidopsis. *Cell* **133,** 523-536. doi: 10.1016/j.cell.2008.03.029.

Liu J, He Y, Amasino R, Chen X (2004) siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in *Arabidopsis*. *Genes & Development* **18,** 2873-2888. doi: 10.1101/gad.1217304

Liu R, How-Kit A, Stammitti L, Teyssier E, Rolin D, Mortain-Bertrand A, Halle S, Liu M, Kong J, Wu C, Degraeve-Guibault C, Chapman NH, Maucourt M, Hodgman TC, Tost J, Bouzayen M, Hong Y, Seymour GB, Giovannoni JJ and Gallusci P (2015) A DEMETER-like DNA demethylase governs tomato fruit ripening. *Proceedings of the National Academy of Sciences of the United States of America* **112,** 10804-10809. doi: 10.1073/pnas.1503362112.

Liu S, Liu Y, Yang X, Tong C, Edwards D, Parkin IAP, Zhao M, Ma J, Yu J, Huang S, Wang X, Wang J, Lu K, Fang Z, Bancroft I, Yang TJ, Hu Q, Wang X, Yue Z, Li H, Yang L, Wu J, Zhou Q, Wang W, King GJ, Pires JC, Lu C, Wu Z, Sampath P, Wang Z, Guo H, Pan S, Yang L, Min J, Zhang D, Jin D, Li W, Belcram H, Tu J, Guan M, Qi C, Du D, Li J, Jiang L, Batley J, Sharpe AG, Park BS, Ruperao P, Cheng F, Waminal NE, Huang Y, Dong C, Wang L, Li J, Hu Z, Zhuang M, Huang Y, Huang J, Shi J, Mei D, Liu J, Lee TH, Wang J, Jin H, Li Z, Li X, Zhang J, Xiao L, Zhou Y, Liu Z, Liu X, Qin R, Tang X, Liu W, Wang Y, Zhang Y, Lee J, Kim HH, Denoeud F, Xu X, Liang X, Hua W, Wang X, Wang J, Chalhoub B, Paterson AH (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploidy genomes. *Nature Communication* **5,** 3930. doi: 10.1038/ncomms4930.

Long Y, Xia W, Li R, Wang J, Shao M, Feng J, King GJ, Meng J (2011) Epigenetic QTL mapping in *Brassica napus*. *Genetics* **189,** 1093-1102. doi: 10.1534/genetics.111.131615.

Lv H, Fang Z, Yang L, Zhang Y, Wang Q, Liu Y, Zhuang M, Yang Y, Xie B, Liu B, Liu J, Kang J, Wang X (2014) Mapping and analysis of a novel candidate Fusarium wilt resistance gene *FOC1* in *Brassica oleracea*. *BMC Genomics* **15,** 1094. doi: 10.1186/1471-2164-15-1094.

Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C, Bendahmane A (2009) A transposon-induced epigenetic change leads to sex determination in melon. *Nature* **461,** 1135-1138. doi: 10.1038/nature08498.

Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nature Reviews Genetics* **15,** 394-408. doi: 10.1038/nrg3683.

Miyaji N, Shimizu M, Miyazaki J, Osabe K, Sato M, Ebe Y, Takada S, Kaji M, Dennis ES, Fujimoto R, Okazaki K (2017) Comparison of transcriptome profiles by *Fusarium oxysporum* inoculation between Fusarium yellows resistant and susceptible lines in *Brassica rapa* L. *Plant Cell Reports* in press. doi: 10.1007/s00299-017-2198-9.

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8,** 4321-4325. doi:10.1093/nar/8.19.4321

Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q, Rohr NA, Rambani A, Burke JM, Udall JA, Egesi C, Schmutz J, Grimwood J, Jackson SA, Springer NM, Schmitz RJ (2016) Widespread natural variation of DNA methylation within angiosperms. *Genome Biology* **17,** 194. doi: 10.1186/s13059-016-1059-0

Osabe K, Sasaki T, Ishikawa R, and Fujimoto R (2012) The role of DNA methylation in plants. In 'DNA Methylation: Principles, Mechanisms and Challenges'. (Eds T.V.T.a.G. Sablok) pp. 35-66 (NOVA publishers).

Parkin IAP, Koh C, Tang H, Robinson SJ, Kagale S, Clarke WE, Town CD, Nixon J, Krishnakumar V, Bidwell SL, Denoeud F, Belcram H, Links MG, Just J, Clarke C, Bender T, Huebert T, Mason AS, Pires JC, Barker G, Moore J, Walley PG, Manoli S, Batley J, Edwards D, Nelson MN, Wang X, Paterson AH, King G, Bancroft I, Chalhoub B, Sharpe AG (2014) Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biology* **15,** R77. doi: 10.1186/gb-2014-15-6-r77.

Sasaki T, Fujimoto R, Kishitani S, Nishio T (2011) Analysis of target sequences of

DDM1s in *Brassica rapa* by MSAP. *Plant Cell Reports* **30,** 81-88. doi: 10.1007/s00299-010-0946-1.

Saze H, Kakutani T (2007) Heritable epigenetic mutation of a transposon-flanked *Arabidopsis* gene due to lack of the chromatin-remodeling factor DDM1. *The EMBO Journal* **26,** 3641-3652. doi: 10.1038/sj.emboj.7601788

Saze H, Sasaki T, Kakutani T (2008) Negative regulation of DNA methylation in plants. *Epigenetics* **3,** 122-124.

Schield DR, Walsh MR, Card DC, Andrew AL, Adams RH, and Castoe TA (2016) EpiRADseq: scalable analysis of genomewide patterns of methylation using next-generation sequencing. *Methods in Ecology and Evolution* **7,** 60-69. doi: 10.1111/2041-210X.12435

Shimizu M, Fujimoto R, Ying H, Pu ZJ, Ebe Y, Kawanabe T, Saeki N, Taylor JM, Kaji M, Dennis ES, Okazaki K (2014) Identification of candidate genes for fusarium yellows resistance in Chinese cabbage by differential expression analysis. *Plant Molecular Biology* **85,** 247-257. doi:10.1007/s11103-014-0182-0

Shimizu M, Pu Z, Kawanabe T, Kitashiba H, Matsumoto S, Ebe Y, Sano M, Funaki T, Fukai E, Fujimoto R, Okazaki K (2015) Map-based cloning of a candidate gene conferring Fusarium yellows resistance in *Brassica oleracea*. *Theoretical and Applied Genetics* **128,** 119–130. doi: 10.1007/s00122-014-2416-6.

Tarutani Y, Shiba H, Iwano M, Kakizaki T, Suzuki G, Watanabe M, Isogai A, Takayama S (2010) *Trans*-acting small RNA determines dominance relationships in *Brassica* self-incompatibility. *Nature* **466,** 983-986. doi: 10.1038/nature09308.

Vaughn MW, Tanurdžić M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, Dedhia N, McCombie WR, Agier N, Bulski A, Colot V, Doerge RW, Martienssen RA (2007) Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biology* **5,** e174. doi: 10.1371/journal.pbio.0050174

Verkest A, Byzova M, Martens C, Willems P, Verwulgen T, Slabbinck B, Rombaut D, Van de Velde J, Vandepoele K, Standaert E, Peeters M, Van Lijsebettens M, Van

Breusegem F, De Block M (2015) Selection for improved energy use efficiency and drought tolerance in canola results in distinct transcriptome and epigenome changes. *Plant Physiology* **168,** 1338-1350. doi: 10.1104/pp.15.00155.

Wang M, Qin L, Xie C, Li W, Yuan J, Kong L, Yu W, Xia G, Liu S (2014) Induced and constitutive DNA methylation in a salinity-tolerant wheat introgression line. *Plant & Cell Physiology* **55,** 1354-1365. doi: 10.1093/pcp/pcu059

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F, Huang S, Li X, Hua W, Wang J, Wang X, Freeling M, Pires JC, Paterson AH, Chalhoub B, Wang B, Hayward A, Sharpe AG, Park BS, Weisshaar B, Liu B, Li B, Liu B, Tong C, Song C, Duran C, Peng C, Geng C, Koh C, Lin C, Edwards D, Mu D, Shen D, Soumpourou E, Li F, Fraser F, Conant G, Lassalle G, King GJ, Bonnema G, Tang H, Wang H, Belcram H, Zhou H, Hirakawa H, Abe H, Guo H, Wang H, Jin H, Parkin IAP, Batley J, Kim JS, Just J, Li J, Xu J, Deng J, Kim JA, Li J, Yu J, Meng J, Wang J, Min J, Poulain J, Wang J, Hatakeyama K, Wu K, Wang L, Fang L, Trick M, Links MG, Zhao M, Jin M, Ramchiary N, Drou N, Berkman PJ, Cai Q, Huang Q, Li R, Tabata S, Cheng S, Zhang S, Zhang S, Huang S, Sato S, Sun S, Kwon SJ, Choi SR, Lee TH, Fan W, Zhao X, Tan X, Xu X, Wang Y, Qiu Y, Yin Y, Li Y, Du Y, Liao Y, Lim Y, Narusaka Y, Wang Y, Wang Z, Li Z, Wang Z, Xiong Z, Zhang Z (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics* **43,** 1035-1039. doi: 10.1038/ng.919.

Zhang X, Shiu SH, Cal A, Borevitz JO (2008) Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays. *PLoS Genetics* **4,** e1000032. doi: 10.1371/journal.pgen.1000032.

Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. *Cell* **126,** 1189-1201. doi: 10.1016/j.cell.2006.08.003

Zhu QH, Shan WX, Ayliffe MA, Wang MB (2016) Epigenetic mechanisms: an

emerging player in plant-microbe interactions. *Molecular Plant-Microbe Interactions* **29,** 187-196. doi: 10.1094/MPMI-08-15-0194-FI.

Figure legends

Figure 1. Comparison of the two replicates for MeDIP-seq of RJKB-T23 and RKKB-T24. The RPKM of each sliding window per 100 kb was compared between replicate 1 (single end, 35bp) and replicate 2 (paired end, 50bp) of MeDIP-seq in RJKB-T23 (top left panel) and RJKB-T24 (top right panel). Correlation coefficient in each window of two replications was 0.86 and 0.87 in RJKB-T23 and RJKB-24, respectively. Graphical representation of distribution of DNA methylation levels (log 10 score of RPKM) in a sliding 100 kb window across chromosomes of RJKB-T23 and RJKB-T24 is shown in supplementary figure 1 and 2

Figure 2. Box plots of the expression levels of log 2 score of FPKM in genes having DNA methylation peaks in the 2 kb upstream regions, exon, intron, or 2 kb downstream regions of RJKB-T23 and RJKB-T24. Total indicates the log 2 score of FPKM in all genes (FPKM < 0.01).

Figure 3. Validation of DNA methylation state by chop-PCR. PCR was performed using genomic DNA digested by *Hpa* II as a template. Eight genes (Bra010682, Bra015165, Bra017403, Bra010590, Bra018542, Bra038263, Bra016440, and Bra037713) were methylated and two genes (Bra001846 and Bra023446) were not methylated. Four independent plants were examined.

Figure 4. Differentially methylated regions between RJKB-T23 and RJKB-T24. DNA methylation peaks were observed only in RJKB-T24 (a) or in RJKB-T23 (b) in the 2 kb upstream regions (upper), exon/intron regions (middle), or 2 kb downstream regions (bottom). Black bar represents 1kb. The boxes of second lane show the interspersed

repeats regions (IRRs)

Figure 5. The number of DNA methylation peaks detected only in the genic region (2 kb upstream, exon, intron, and 2 kb downstream regions) of RJKB-T23 (T23-SME, specifically methylated genes in T23) or RJKB-T24 (T24-SME).

Figure S1. Chromosomal distribution of DNA methylation levels (log 10 score of RPKM) in a 100 kb sliding windown in RJKB-T23.

Figure S2. Chromosomal distribution of DNA methylation levels (log 10 score of RPKM) in a 100 kb sliding windown in RJKB-T24.

Figure S3. Percentage of mapped reads on the interspersed repeats regions (IRRs) using Input-DNA-seq and MeDIP-seq data in RJKB-T23 and RJKB-T24. SE, single-end; PE, paired-end

Figure S4. Proportion of mapped reads of Input-DNA-seq and MeDIP-seq in 2 kb upstream, exon, intron, and 2 kb downstream regions in RJKB-T23 and RJKB-T24.

Figure S5. Venn diagram of genes having DNA methylation peaks in 2 kb upstream, exon, intron, and 2 kb downstream regions of RJKB-T23 compared with RJKB-T24.

Figure S6. Visualization of DNA methylation peaks by Integrative Genomics Viewer (IGV). (a) DNA methylation peaks were observed in interspersed repeats regions (IRRs) of intergenic regions. (b) DNA methylation peaks were observed in intron region of Bra033012 overlapped with IRRs.

Figure S7. Classification into seven groups of expression levels of genes having

MeDIP-peaks in the 2 kb upstream (Up (2k)), exon, intron, and 2 kb downstream regions (Down (2k)). Group-0, No mapped read; Group-1, $\log_2$ (FPKM)<-3.0; Group-2, -3.0=< $\log_2$ (FPKM)<0.0; Group-3, 0.0=< $\log_2$ (FPKM)<3.0; Group-4, 3.0=< $\log_2$ (FPKM)<6.0; Group-5, 6.0=< $\log_2$ (FPKM)<9.0; Group-6, 9.0=< $\log_2$ (FPKM).

Figure S8. Differentially methylated regions between RJKB-T23 and RJKB-T24. DNA methylation levels in RJKB-T23 were higher (a) or lower (b) than in RJKB-T24 in the genic regions (200 bp upstream/exon/intron/ 200bp downstream). Black bar represents 500 bp.