

MiRNA-mediated Changes in DNA Methylation Affect the Expression of Genes Involved in the Thickness of Pod Canopy Trait in Brassica Napus

Zhiyou Chen

Southwest University

Ledong Jia

Southwest University

Yuanyuan Wan

Southwest University

Jinqi Ma

Southwest University

Kun Lu

Southwest University

Cunmin Qu

Southwest University

Jiana Li (✉ ljn1950@swu.edu.cn)

Southwest University

Research Article

Keywords: DNA methylation, thickness of pod canopy, Brassica napus, nitrogen-related pathway

Posted Date: February 4th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-136648/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Methylation plays an important role in regulating crop development, but little is known about how methylation regulates plant architecture in rapeseed (*Brassica napus*). Here, we examined how methylation affects the TPC (thickness of pod canopy) trait in rapeseed by performing genome-wide methylation analysis of two extreme TPC lines.

Results: We detected significant differences in overall methylation levels between the high- and low-TPC lines in the CG, CHG, and CHH contexts in the promoters of genes in the stem apex and flower bud. In flower buds, 26 genes had significantly higher methylation levels in the high-TPC samples compared to the low-TPC samples, resulting in significantly reduced gene expression. By contrast, in the stem apex samples, the promoter regions of 22 genes were hypermethylated in the high- vs. low-TPC samples. The promoters of 19 and 21 genes had significantly reduced methylation levels in the flower bud and stem apex, respectively, of the high- vs. low-TPC samples, resulting in significantly higher expression levels. Some of these differentially expressed genes are associated with TPC-related traits, such as BnaC03g53050D (UBC32), BnaA05g26660D (CYSB), BnaA10g07880D (TCP 1), BnaAnng09670D (SMP1), BnaA09g02000D (SDH2-2), BnaC01g12960D (NRT1.8), and BnaC09g30490D (TAF15b). In addition, 14 important genes related to growth and development were differentially regulated between the two groups due to miRNA-mediated differences in methylation levels in their promoters. For example, hypermethylation in the promoter region of BnaCnng64040D (Lipase family protein) mediated by miR159a led to significantly reduced gene expression in flower buds of high-TPC vs. low-TPC lines.

Conclusions: These results, together with our previously generated RNA-seq and miRNA profiling data, indicate that both methylation and miRNAs are involved in regulating the expression of genes in nitrogen-related metabolic pathways, thereby affecting the TPC trait in *B. napus*, providing a reference for uncovering the molecular mechanism regulating this crucial trait.

1. Introduction

Oilseed rape (*Brassica napus*, $2n = 38$, AACC) is one of the most important oilseed crops worldwide. Yield improvement is an extremely important goal for all rapeseed breeders. Modifying plant type is the most important way to increase rapeseed yields. Thickness of pod canopy (TPC), i.e., the thickness of the plant canopy from the bottom-most effective (seed-bearing) pod to the uppermost effective pod, is a key trait that determines the three-dimensional structure of plants. TPC is directly related to other important traits, such as economic yield (EY, grain weight per plant), plant height (PH), pod terminal height (PTH), stem height (SH), first effective branch height (FEBH), height of the lowest pod (HLP), first effective branch number (FEBN), main inflorescence effective length (MIEL), and first ineffective branch number (FUBN) [1]. Hence, understanding the molecular mechanism of TPC could shed light on the formation of plant architecture and help breeders further increase production in *B. napus*.

In recent years, many studies have focused on traits in crops. PH is an important factor affecting rice yield. OsMPH1 improves grain yield by regulating PH in rice [2]. MiR319 expression induces dwarfism to suppress PH in rice [3]. A genome-wide association study (GWAS) of PH and primary branch number in rapeseed revealed eight PH-related quantitative trait loci (QTLs) on chromosomes A03, A05, A07, and C07 and five PB-related QTLs on chromosomes A01, A03, A07, and C07 [4]. GWAS identified four, five, and seven SNPs for FEBH, FEBN, and PH, respectively, in *B. napus* and uncovered many genes associated with these traits [5]. We recently reported that TPC is regulated by miRNAs and identified many candidate genes for TPC based on RNA-Seq and miRNA profiling analyses [1]. Many genes involved in nitrogen-related responses were dramatically differentially expressed in high- vs. low-TPC lines, such as *ASP5*, *ASP2*, *ASN3*, *ATCYSC1*, *PAL2*, *APT2*, *CRTISO*, and *COX15* [1], suggesting that nitrogen metabolism plays important roles in regulating TPC in *B. napus*.

DNA methylation is a critical epigenetic modification in many plants. The methylation and demethylation of genes are dynamically regulated during plant growth [6]. Plant DNA is methylated in three sequence contexts: symmetrical mCG and mCHG and asymmetrical mCHH (H = A, T, or C) [7–8]. Methylation is regulated and maintained by independent pathways that can be broadly classified into maintenance of methylation and *de novo* methylation pathways [6]. Due to their symmetrical nature, CG and CHG sites are substrates for maintenance methyltransferases, which recognize hemimethylated DNA after replication and methylate newly synthesized unmethylated strands [9]. *De novo* methyltransferases establish DNA methylation across all sequence contexts and are required for maintaining asymmetric (CHH) methylation.

Small RNAs and proteins related to modified histones affect DNA methylation [6]. DNA methylation and demethylation are distinct processes: the latter is regulated by 5-methylcytosine DNA glycosylase enzymes in plants [10]. 5-methylcytosine is removed from DNA by base excision repair, which is essential for the expression of imprinted genes and endosperm development [11].

In plants, DNA methylation plays a key role in the activation of maturation-inducing genes and the inhibition of maturation-inhibiting genes in tomato (*Solanum lycopersicum*) [12]. DNA methylation increases substantially during citrus (*Citrus* sp.) fruit development and maturation [13]. DNA methylation is closely related to the development of many grain crops. Genome-wide changes in DNA methylation are related to drought stress tolerance [14] and cadmium stress in plants [15]. Overall DNA methylation levels increase during somatic embryogenesis in soybean (*Glycine max*) [16]. Cytosine methylation plays a positive role in regulating isoflavone synthase gene expression and isoflavonoid biosynthesis in soybean seeds [17]. Low zinc levels lead to the loss of methylation in maize (*Zea mays*) roots; conversely, the loss of methylation in root cells leads to zinc deficiency [18]. DNA methylation changes dynamically in response to heat stress in maize seedlings [19]. Lead, cadmium, and zinc toxicity alter DNA methylation levels, thereby enhancing heavy metal tolerance in wheat (*Triticum aestivum*) [20]. Genome-wide DNA methylation profiling of flower

buds revealed the role of DNA methylation in the molecular regulation of genic male sterility in *B. napus* [21]. Finally, short-term heat-shock treatment of cultured *B. napus* microspores led to global changes in DNA methylation, indicating that DNA methylation plays a key role in heat-stress responses [22].

Methylation also affects plant architecture. The proper H3K4me3 levels, which are regulated by COMPASS-like complexes, are critical for rice development and can affect flowering and branching [23]. In addition, changes in methylation at the rice (*Oryza sativa*, Os) *FERTILIZATION INDEPENDENT ENDOSPERM 2* locus play important roles in regulating plant height and yield in rice [24]. Changes in methylation are also related to final plant height in *Arabidopsis thaliana* [25]. However, our understanding of the mechanisms underlying how methylation regulates plant-type traits in crops remains limited.

Small RNA-directed DNA methylation (RdDM) is a key regulatory pathway that affects numerous plant traits by altering the expression of various genes. The ARGONAUTE 4 (AGO4)-dependent RdDM pathway represses the expression of *HOMOLOG OF RPW8 4* and alters the response to submergence in Arabidopsis (*Arabidopsis thaliana*) [26]. RdDM inhibits the expression of *APETALA3 (AP3)*; transgenic Arabidopsis plants harboring promoter lines showed abnormal stamens and petals [27]. The RdDM pathway is also involved in regulating the activity of the poly(A) polymerase PAPS1 in Arabidopsis, thereby affecting sporophyte and pollen development [28]. RdDM also regulates seed dormancy in plants [29]. This pathway regulates the expression of numerous genes in crops and affects the corresponding traits. ITRAQ (isobaric tags for relative and absolute quantitation) analysis of the leaf proteome indicated that the RdDM pathway plays a key role in defense against geminivirus-betasatellite infection in tobacco (*Nicotiana tabacum*) [30]. The reinforcement of DNA methylation at CHH sites regulated by the RdDM pathway leads to decreased gene expression, thereby influencing somatic embryogenesis in soybean [16]. However, the role of the RdDM pathway in regulating plant-type traits in crops has not been reported.

TPC is an important plant-type trait that is closely associated with various phenotypes. We recently identified many genes related to the formation of TPC and identified miRNAs that limit their expression in *B. napus* [1]. In the current study, we performed genome-wide methylation analysis of high- and low-TPC *B. napus* lines. The results of this study, combined with previously obtained RNA- and sRNA-sequencing data, shed light on the molecular mechanism of TPC in this important oilseed crop.

2. Materials And Methods

2.1. Plant Materials and Bisulfite Sequencing

We performed methylation-sequencing analysis of plants cultivated at the Chongqing Rapeseed Engineering Research Center, Southwest University, Chongqing, China (106.40°E, 29.80°N). *B. napus* lines YC4 (SWU71) and YC33 (10-1047) showed relatively high TPC values based on three years of continuous observation, while YC11 (Zhongshuang11) and YC15 (Zhongyou821) showed relatively low TPC values. We therefore used these lines for comparative analyses. When the plants bolted and produced tiny flower buds, tissue samples from the stem apex and nearby flower buds were collected: this represents the key stage for the formation of TPC (after this stage, the stem begins to elongate and pollen begins to develop). We collected at least five sets of samples per line, each set from a different individual. Samples were collected at approximately the same time, and all samples from each line were mixed into one pool. Total DNA was isolated from the samples using a Rapid Plant Genomic DNA Isolation Kit (Sangon, China) according to the manufacturer's protocols.

After checking sample quality, we added lambda DNA to the sample as a negative control and used a Covaris S220 sonicator to randomly break the genomic DNA into 200–300 bp fragments. We performed end repair on the broken DNA fragments, added an A tail, and attached them to sequencing adapters whose cytosines were modified by methylation. We performed bisulfite treatment (using an EZ DNA Methylation Gold Kit, Zymo Research), which converted unmethylated C to U (converted to T after PCR amplification), whereas methylated C remained unchanged. PCR was performed to obtain the final DNA library. A library construction flowchart is shown in Supplementary Fig. 1.

2.2. Analysis of Sequencing Data

After obtaining the original sequencing reads, we performed bioinformatic analysis as shown in Supplementary Fig. 2. We preprocessed the reads (in FASTQ format) produced by Illumina sequencing using Trimmomatic (version 0.35). The steps were as follows: First, reads containing adapter sequences were filtered out; Second, reads with N (unknown bases) < 10% were deleted; third, reads containing > 50% low-quality bases (PHRED score \leq 20) were removed. The Q20, Q30, and GC contents of the data were calculated at the same time. The remaining reads that passed all the filtering steps were counted as clean reads and used for subsequent analysis.

2.3 Aligning the Reads to the Reference Genome

Bismark software (version 0.12.5) was used to compare the bisulfite-treated reads to the reference genome using default values. The reference genome was converted to a bisulfite-converted version (C-T and G-A conversion) and indexed by Bowtie2 (version 2.2.5). The reads were also converted into bisulfite-converted versions (C-T and G-A conversion) and directly compared with the converted reference genome. The reads were then compared with the normal genomic sequence to infer the methylation status of each cytosine position in the reads. Reads pairs with the same coordinates in the genome were treated as duplicates and deleted before the methylation status was determined to avoid potential calculation bias

of the methylation level. The non-conversion rate of bisulfite was calculated as the percentage of sequenced cytosine to the reference cytosine in the lambda genome. The basic concept behind Bismark comparison is shown in Supplementary Fig. 3.

2.4 Analysis of Differentially Methylated Regions

Differentially methylated regions (DMRs) were identified using swDMR software (<https://sourceforge.net/projects/swdmr/>) with the conditions read coverage ≥ 5 , methylation level difference ≥ 0.1 or fold change ≥ 2 , corrected p-value < 0.01 using the sliding window method. The window size was set to 1000 bp and the step size was 100 bp. Fisher's exact test was used to detect DMRs. Genes whose functional regions overlapped with DMRs by at least 1 bp were defined as DMR-associated genes (DMGs).

2.5 GO Analysis of DMGs

The DMGs, which were corrected for gene length deviations, were subjected to gene ontology (GO) analysis using the Goseq R package. GO terms with corrected p-values < 0.05 were considered to be significantly enriched by the DMGs. All DMGs were annotated with BGI Web Gene Ontology Annotation plot (<http://wego.genomics.org.cn/>).

2.6 Data Verification Using Traditional Bisulfite Sequencing PCR

Traditional bisulfite sequencing PCR was used to verify the WGBS data. Each 1 μg sample of genomic DNA was subjected to sodium bisulfite using an EpiTect Bisulfite kit according to the manufacturer's instructions (Qiagen). The processed DNA was purified with a Qiagen PCR purification kit (Cat. No. 28106) and used as a template for bisulfite sequencing PCR (BS-PCR). The primers were designed using Kismeth (http://katahdin.mssm.edu/kismeth/primer_design.pl). 2 \times Taq Master Mix (Vazyme Biotech, Nanjing, China) was used for BS-PCR. The amplicon was cloned into pMD18-T (Takara Biotechnology, Dalian, China), and 10–14 positive clones per PCR product were sequenced by Biotechnology (Shanghai) Co., Ltd. The sequencing results were processed using the Seqman program of the DNASTAR software package to remove the vector and primer sequences and analyzed on the Kismeth website [31]. BnaIND.a-A3 was used to determine the conversion efficiency of bisulfite [32].

3. Results

3.1. FastQC and Trimming of Raw Data

With the rapid development of high-throughput sequencing technology, massive amounts of data can now be obtained by sequencing. Bioinformatic analysis is used to obtain useful information from these data. Quality control (QC) is the first step in data analysis. FastQC is commonly used to assess data quality. We used FastQC to perform basic statistical analysis of the quality of the raw reads and the results are shown in Fig. 1.

The main purpose of data filtering is to remove low-quality data and ensure the quality of clean data. We trimmed off the sequencing adapters and low-quality fragments from the raw sequencing data using Trimmomatic software and the results are shown in Table 1.

Table 1
Raw data quality control statistics

LibraryID	Sample name	Raw_reads	Raw_bases(G)	Clean _reads	Clean _bases(G)	Clean _ratio(%)	Q20(%)	Q30(%)	GC(%)	BS conversion rate(%)
YC4F	YC4F	124760799	37.43	122980334	35.07	93.69	96.87	90.99	22.44	99.832
YC33F	YC33F	105409809	31.62	103965545	29.77	94.15	97.06	91.39	22.65	99.847
YC11F	YC11F	103627338	31.09	102281721	29.23	94.02	97.12	91.55	21.98	99.838
YC15F	YC15F	114784401	34.44	113374092	32.43	94.16	97.20	91.95	23.04	99.860
YC4S	YC4S	119810866	35.94	118187983	33.95	94.46	97.36	92.26	22.32	99.867
YC33S	YC33S	123763449	37.13	122165163	35.00	94.26	97.25	92.02	22.68	99.852
YC11S	YC11S	119497670	35.85	118197197	33.78	94.23	97.30	91.89	22.82	99.856
YC15S	YC15S	129328290	38.80	127744829	36.50	94.07	97.21	91.66	21.81	99.823

(1) LibraryID: library name; (2) Sample name: sample name; (3) Raw Reads: statistical raw sequence data, counting the number of sequencing sequences per file in four units; (4) Raw Bases (G): the number of Raw Reads multiplied by the length of the sequencing sequence, in units of G; (5) Clean Reads: The calculation method is the same as Raw Reads, except that the statistical file is the sequencing data filtered by Trimming, and the subsequent bioinformatics analysis is based on Clean Reads; (6) Clean Bases (G): Bases included in Clean Reads Number, in G; (7) Clean Ratio (%): refers to the proportion of Clean Bases in Raw Bases; (8) Q20 (%): refers to the proportion of bases with Q values greater than 20 in all bases; Q30 (%): refers to the proportion of bases with Q values greater than 30 in all bases; (10) GC Content (%): the sum of the number of bases G and C as a percentage of the total number of bases; (11) BS Conversion rate (%): refers to the ratio of Bisulfite converting C to T.

3.2 Analysis of Clean Reads

A comparison of the clean reads with the reference genome showed that the number of unique reads accounted for ~ 50% of the total, and the ratio of repetitive sequences in the reads to total reads was ~ 10% (Table 2).

Table 2
Comparison list of reads and reference genome

Samples	Total reads	Mapped reads	Mapping rate(%)	Duplication rate(%)
YC4F	122980334	67848250	55.17	11.47
YC33F	103965545	58366256	56.14	12.40
YC11F	102281721	57707346	56.42	11.48
YC15F	113374092	54748349	48.29	13.96
YC4S	118187983	63738779	53.93	14.39
YC33S	122165163	67630634	55.36	14.55
YC11S	118197197	65008458	55.00	13.86
YC15S	127744829	71217742	55.75	13.20

(1) Samples: sample names; (2) Total reads: the number of sequencing sequences (Clean data) after processing; (3) Mapped reads: the number of reads that are uniquely aligned to the reference genome; (4) Mapping rate (%): the ratio of the number of reads on the unique alignment to the total number of reads; (5) Duplication rate (%): the proportion of repeats in the reads to total reads.

To fully detect the methylation of a genome, the NIH Roadmap Epigenomics Project (<http://www.roadmapepigenomics.org/protocols>) recommends a sequencing depth of 30× or more. Therefore, we used a sequencing depth of 30× to perform methylation sequencing of all samples and to calculate the coverage of each single-base site in the genome (that is, the number of reads that support that site). The coverage statistics and distribution diagrams are shown in Table 3 and Fig. 2, respectively.

Table 3
Statistics list of genomic coverage

Samples	sites_num	sites_covgMean	sites_numCovg1	sites_numCovg5	sites_numCovg10
YC4F	627510559	20.23	73.80	68.18	63.60
YC33F	618507269	17.25	72.74	67.00	61.77
YC11F	627161592	17.23	73.76	68.22	62.04
YC15F	608705698	15.91	71.59	63.52	56.27
YC4S	628148992	18.42	73.87	68.38	63.57
YC33S	622392432	19.51	73.20	67.81	63.47
YC11S	630245667	18.87	74.12	68.56	63.06
YC15S	633955599	20.82	74.56	69.36	65.49

(1) Samples: sample name; (2) sites_num: number of loci detected on the genome; (3) sites_covgMean: average coverage depth of all loci in the genome; (3) sites_numCovg1: greater than or equal to 1× on the genome The number of bases accounted for the total length of the genome; (4) sites_numCovg5: the number of bases on the genome greater than or equal to 5 × the depth of sequencing accounted for the total length of the genome; (5) sites_numCovg10: the number of bases on the genome greater than or equal to 10 × the depth of sequencing accounted for the total length of the genome proportion.

The coverage level of the C site is an important indicator of the sequencing depth in the methylation test. We calculated the coverage of the C site separately, as well as the coverage of the C site under each context (CpG, CHH, CHG) (that is, the number of reads that support that context). The results and cumulative distribution diagram are shown in Table 4 and Fig. 3, respectively. Among the three methylation environments (CG, CHG, and CHH), CG had the highest coverage (47.99–51.59%), followed by CHG (18.73–24.75%) and CHH (2.34–4.59%).

Table 4
Statistics on methylation status of C site

sample	C_covgMean	C(Mb)	CG(Mb)	CHG(Mb)	CHH(Mb)	MeanC(%)	MeanCG(%)	MeanCHG(%)	MeanCHH(%)
YC4F	8.8	2336.7	344.7	365.8	1626.2	11.89	48.70	19.67	2.34
YC33F	7.1	1905.2	280.7	295.8	1328.7	13.24	51.59	23.92	2.76
YC11F	7.3	1959.0	281.7	298.9	1378.5	11.92	49.62	20.48	2.37
YC15F	7.4	1967.7	305.6	319.7	1342.5	13.00	50.90	18.73	3.00
YC4S	7.7	2063.7	302.6	320.7	1440.4	13.22	49.55	20.80	3.91
YC33S	8.3	2201.7	320.5	339.5	1541.7	14.34	50.22	24.75	4.59
YC11S	8.1	2163.7	321.5	339.5	1502.7	13.67	49.29	21.84	4.20
YC15S	8.9	2372.5	336.7	359.1	1676.7	12.40	47.99	18.36	3.97

(1) sample: sample name; (2) C_covgMean: average coverage depth of all C sites in the genome; (3) C(Mb): the number of bases aligned to the position of the genomic C; (5) CG (Mb) : the number of bases on the cytosine in the genomic region of the genomic region; (6) CHG (Mb): the number of bases on the cytosine in the CHG region of the genomic region; (7) CHH(Mb): alignment The number of bases on the cytosine in the CHH region of the genome; (8) MeanC (%): the average methylation level of all C sites in the genome; (9) MeanCG (%): the average methylation level of the CG region; (10) MeanCHG (%): average methylation level in the CHG region; (11) MeanCHH (%): average methylation level in the CHH region.

3.3 Types and Distribution of Methylation in the high- and low-TPC lines

We performed statistical analysis of the methylated C sites in three sequence environments (CG, CHH, CHG, where H stands for A, C, or T) in the two plant parts (young buds and stem apex) for the two sets of samples: high-TPC (YC4, YC33) and low-TPC (YC11, YC15) lines (Table 5 and Fig. 4). The percentage of methylated C sites in the three sequence regions of all samples represents the percentage of the total number of C sites in that region. The percentages of CG sites were the highest (33.32–37.32%), followed by CHG (21.05–26.69%) and CHH (2.88–6.52%). The overall average level of methylation of the three sequences in all samples (from high to low) was CG (43.99%), CHG (29.09%), and CHH (26.92%).

Table 5
Summary of Whole Genome Methylation

Samples	mC percent(%)	mCpG percent(%)	mCHG percent(%)	mCHH percent(%)
YC4F	11.31%	37.15%	23.92%	3.42%
YC33F	10.61%	33.59%	23.43%	3.26%
YC11F	9.96%	33.32%	21.05%	2.88%
YC15F	10.88%	35.98%	21.55%	3.54%
YC4S	12.1%	34.88%	23.29%	5.12%
YC33S	13.78%	36.2%	26.69%	6.52%
YC11S	12.8%	35.64%	24.51%	5.71%
YC15S	13.11%	37.32%	22.72%	6.17%

(1) Samples: sample name; (2) mC percent (%): the percentage of methylation C sites in the whole genome as a percentage of all C sites; (3) mCpG percent (%): A in the CG context region The number of base C sites in the total number of C sites in the region; (4) mCHG percent (%): the percentage of methylated C sites in the CHG context region as a percentage of the total number of C sites in the region; (5) mCHH percent (%): The percentage of methylated C sites in the CHH context region as a percentage of the total number of C sites in the region.

The average methylation levels of mC (10.96%), mCpG (35.37%), mCHG (23.68%), and mCHH (3.34%) sites in young buds were higher in the high-TPC lines than in the low-TPC lines (10.42%, 34.65%, 21.30% and 3.21%, respectively). In the stem apex, although the methylation level at the mCHG sites was higher in the high-TPC lines (24.99%) than in the low-TPC lines (23.62%), the methylation levels at the mC, mCpG, and mCHH sites were lower in the high-TPC lines (12.94%, 35.54%, and 5.82%, respectively) than in the low-TPC lines (12.96%, 36.48%, and 5.94%, respectively). A comparison of the average methylation levels of the different plant parts in all samples showed that the methylation levels at mC, mCpG, mCHG, and mCHH sites were lower in young buds (10.69%, 35.01%, 22.49%, and 3.28%, respectively) than in the stem apex (12.95%, 36.01%, 24.30%, and 5.88%, respectively). An analysis of the methylation level of each chromosome in each sample showed that the ratios of methylated C sites to total methylated C sites on the chromosome in different sequence environments were consistent across all materials: CG had the highest ratio, followed by CHG and CHH (Fig. 5).

3.4 Methylation Density on Chromosomes and Distribution of Methylation Levels in Genes

B. napus is a heterotetraploid with two sets of chromosomes (A and C). Circos diagrams are commonly used to show the distribution of methylation density on chromosomes [33–34]. There were significant differences in the distribution of methylation density on C vs. A chromosomes in all samples. Even though A chromosomes contain more genes than C chromosomes, the methylation density of C chromosomes was generally higher than that of A chromosomes (Fig. 6).

We analyzed the average methylation levels of C sites in the CG, CHG, and CHH contexts in various functional genomic regions (such as promoter, exon, intron, 5' UTR, 3' UTR, and so on). Here, the 2-kb region upstream of the transcription start site (TSS) was considered to be the promoter region. The distribution of the average methylation levels in the functional elements of the genes is shown in Fig. 7. In both the high- and low-TPC lines, the results were consistent in all contexts. The promoter regions had the highest methylation levels, followed by introns. We also analyzed the methylation levels in the upstream and downstream regions of each gene in each sample. Specifically, we calculated the average methylation level of the C site in each gene body, 2 kb upstream of the TSS, and 2 kb downstream of the transcription termination site in each context; the results are shown in Supplementary Fig. 4. The average methylation levels were higher in the regions 2 kb upstream and downstream of the TSS than in the gene body regions in all samples.

3.5 Comparison of the Overall Methylation Levels of High- and Low-TPC Lines

We constructed a Circos diagram to display differences in the methylation levels in different samples from the high- vs. low-TPC lines [33, 35]. We detected significant differences in the overall methylation levels of multiple chromosomes between the two groups of lines (Supplementary Fig. 5A-C). An analysis of the differences in methylation levels in different gene functional regions in the high- and low-TPC lines revealed significant differences in the average methylation levels of C sites under all three contexts (CG, CHG, and CHH) between groups, especially in the promoter region (2 kb upstream of the TSS) (Supplementary Fig. 5D-F). We compared the methylation levels of the 2 kb upstream and downstream regions of genes between the high- and low-TPC lines at C sites under all three contexts and detected significant differences in both stems and flower buds (Supplementary Fig. 5G-I).

We also examined the differences in overall methylation levels between stem and flower buds in both groups. The C methylation level of CG showed the greatest difference between stems and young buds, followed by CHG. By contrast, for all samples, there was no significant difference in the level of C methylation in the CHH context between stems and young buds (Fig. 8).

3.6. DMR Analysis of High- vs. Low-TPC lines

We used DSS software to identify DMRs between the high- and low-TPC lines [36–38]. We generated a Circos diagram to visualize the distribution of DMRs in the genome and the results of mapping analysis. We detected numerous DMRs in all three sequence contexts (CG, CHG, and CHH) between the two sets of lines in both the stem apex and flower buds (Supplementary Fig. 6). All comparisons generated consistent results. The most significant DMRs included both hypermethylated and hypomethylated regions in the CG sequence context (Supplementary Fig. 6).

We performed statistical mapping of the DMR anchoring areas (such as promoter, exon, intron, CGI, CGI shore, repeat, TSS, and TES, regions) to distinguish hypermethylated vs. hypomethylated DMRs (Fig. 9). We detected both hypermethylated and hypomethylated DMRs in all functional elements in both the high- and low-TPC lines in the stem apex and flower bud. Notably, there were many DMRs in the promoter regions of all samples, including the CG, CHG, and CHH contexts.

3.7 DMR analysis of the promoters of high- vs. low-TPC lines

Methylation of the promoters of plant genes plays an important role in regulating gene expression, thus affecting plant traits [39]. Therefore, we identified genes with significantly different methylation levels in their promoter regions. In the stem apex, 427 genes were hypermethylated and 547 genes were hypomethylated in the promoter regions in the high-TPC lines compared to the low-TPC lines (Fig. 10C and D). In flower buds, 367 genes were hypermethylated and 531 genes were hypomethylated in the promoter regions of the high-TPC lines (Fig. 10A and B). Compared to the low-TPC lines, 293 genes were hypermethylated and 398 genes were hypomethylated in the promoter regions in both the stem apex and flower buds (Fig. 10E). In addition, for a few genes, the promoter regions were significantly hypermethylated or hypomethylated (Fig. 10E).

3.8 DMRs in Promoters Underlie Differences in Gene Expression Between High- and Low-TPC Lines

Increased methylation in the promoter region of a gene inhibits gene expression, while demethylation in the promoter region of a gene promotes gene expression [39]. We combined the current results with our previously reported transcriptome data [1] and identified 26 genes in the high-TPC materials that were significantly hypermethylated in their promoter regions compared to the low-TPC materials in *B. napus* flower buds, resulting in significant decreases in their expression (Fig. 11A and Supplementary Table 1). We also identified 22 genes that were hypermethylated in their promoter regions in the high- vs. low-TPC materials in the stem apex, resulting in significant decreases in their expression (Fig. 11A and Supplementary Table 1). Finally, the methylation levels of the promoters of 19 and 21 genes were significantly reduced in the flower bud and stem apex, respectively, in high- vs. low-TPC materials, which led to significant increases in their expression (Fig. 11B and Supplementary Table 1).

Therefore, the promoters of many genes are affected by hypermethylation or demethylation in the high- vs. low-TPC materials, leading to significant differences in gene expression. Some of these genes are related to the TPC trait. For example, in both the stem apex and flower buds of the high-TPC

lines, the promoter regions of BnaC03g53050D (UBC32), BnaA05g26660D (CYSB), and BnaA10g07880D (TCP1) had low methylation levels in the high-TPC lines, and they were expressed at significantly higher levels in the high-TPC vs. low TPC-lines. The promoters of BnaC03g70780D (auxin associated family protein), BnaC05g36710D (myb family transcription factor), BnaAnng09670D (SMP1), BnaA09g02000D (SDH2-2), and BnaC01g12960D (NRT1.8) were hypermethylated in the high- vs. low-TPC lines, leading to significantly lower expression levels in both buds and the stem apex. The promoter region of BnaC09g30490D (TAF15b) was hypomethylated in the flower buds of the high- vs. low-TPC lines, leading to higher expression in this tissue.

3.9 GO and KEGG Analysis of Differentially Methylated Genes

To investigate the functions of genes whose promoter regions were differentially regulated by methylation and were differentially expressed between the high- and low-TPC lines, we performed GO and KEGG analysis of these genes (Fig. 12). The significantly enriched GO terms are involved in nitrogen metabolism, biosynthesis regulation, cytoskeleton composition, and so on (Fig. 12A). KEGG analysis showed that the significantly enriched pathways mainly involve energy-related processes such as carbon metabolism (Fig. 12B).

3.10 MiRNAs and Methylation Jointly Regulate the TPC Trait

To examine whether miRNAs and methylation are both involved in regulating the expression of genes involved in the TPC trait, we performed a joint analysis of transcriptome, methylation, and miRNA data for the high- vs. low-TPC samples. The promoter regions of 14 important genes related to growth and development were differentially methylated and their expression regulated by miRNAs, resulting in significant differences in gene expression between the two groups (Table 6). For example, the expression of BnaCnng64040D (Lipase family protein) is regulated by miR159a, and hypermethylation in its promoter region led to a significant decrease in gene expression in the flower buds of high- vs. low-TPC samples. The expression of BnaC09g30490D (TAF15b) and BnaC03g09180D is regulated by miR167c and miR827, respectively, and the loss of methylation in their promoter regions resulted in a significant increase in their expression in the flower buds of the high-TPC lines. Both BnaC02g22120D (Spc97/Spc98 family of spindle pole body component) and BnaA07g07800D are regulated by miR319. Increased promoter methylation led to a significant decrease in the expression of these genes in the stem apex of the high-TPC lines. BnaA04g14930D (S-locus lectin protein kinase family protein) and BnaCnng56050D (Pyridoxal phosphate-dependent Transferases superfamily protein) are regulated by miR159, miR319, miR122, and so on. The loss of methylation in the promoters of these genes led to a significant increase in gene expression in the stem apex and flower buds of the high-TPC lines.

Table 6
MiRNAs and methylation are involved in the regulation of gene expression TPC trait

Gene_id	MiRNAs	Methylation	MiRNAs_expression	Gene_expression	Sites	Comparison with Arabidopsis genes
BnaAnng30250D	MIR319a	hypermethylation	up	down	common	-
BnaC04g55380D	MIR9409/MIR6029	hypermethylation	up	down	common	microtubule-associated protein 65 – 4 (MAP65-4)
BnaC08g03460D	MIR5726/MIR9409	hypermethylation	up	down	common	FUNCTIONS IN: sequence-specific DNA binding transcription factor activity
BnaCnng50740D	MIR9563a	hypermethylation	up	down	common	RNA-directed DNA polymerase (reverse transcriptase)-related family protein
BnaA09g11170D	MIR9409	hypermethylation	up	down	flower bud	pentatricopeptide (PPR) repeat-containing protein
BnaCnng64040D	MIR159b	hypermethylation	up	down	common	Lipase family protein
BnaA07g07800D	MIR319/MIR319a	hypermethylation	up	down	stem apex	-
BnaC02g22110D	MIR9409	hypermethylation	up	down	stem apex	-
BnaC02g22120D	MIR319a	hypermethylation	up	down	common	Spc97 / Spc98 family of spindle pole body (SBP) component
BnaA04g14930D	MIR159/MIR319/MIR5726	hypomethylation	down	up	common	S-locus lectin protein kinase family protein
BnaCnng56050D	mir-122/MIR5726/MIR9410	hypomethylation	down	up	common	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein
BnaA06g09210D	MIR395d/MIR827	hypomethylation	down	up	flower bud	IQ-domain 28 (IQD28)
BnaC03g09180D	MIR827	hypomethylation	down	up	flower bud	-
BnaC09g30490D	MIR167c	hypomethylation	down	up	flower bud	TBP-associated factor 15B (TAF15b)

Note: In the “Sites” column, “common” indicates significant differential expression in shoot tips and flower buds. “flower bud” indicates significant differential expression in flower buds, and “stem apex” indicates significant differential expression in shoot tips.

3.11 Verification of the methylation sequencing results

The correlation of methylation levels between samples is an important indicator of the reliability of an experiment and whether the sample selection is reasonable. The closer the correlation coefficient is to 1, the higher the similarity of the methylation patterns between samples. We used the 2-kb/bin sub-sequence environment to calculate the methylation level in each bin and performed Pearson correlation analysis of the data [40]. The correlations coefficients of the CG, CHG, and CHH methylation levels between the stem apex and flower buds in the four high- and low-TPC lines were

very high (Fig. 13). This finding indicates that the overall differences in methylation between the two plants parts were not significant and that the sequencing results are reliable.

To further verify the reliability of the sequencing results, we used a DNA Bisulfite Conversion Kit (Tiangen, Beijing) to process the remaining samples from the sequencing experiment (we randomly selected YC4F and YC4S). We amplified segments of the differentially methylated regions of 20 gene promoters by PCR, followed by sequencing after TA cloning. We compared the sequencing results with the previously generated high-throughput sequencing data. The promoter regions of all 20 genes showed the same methylation patterns in both sets of data (Supplementary Table 2). Therefore, the results were reliable and could be used for further analysis.

4. Discussion

Cytosine DNA methylation is a chemical modification that produces 5-mC, also known as the fifth base of DNA. In plants, DNA methylation occurs in three different sequence environments: symmetric CG and CHG and asymmetric CHH, where H stands for C, A, or T. These DNA methylation patterns are stably inherited through cell division. Changes in DNA methylation can occur spontaneously or be induced by genetic factors and environmental stimuli. High-throughput DNA sequencing is used for single-base resolution transcriptome and epigenome analyses, as well as genome resequencing [41]. The integration of these omics-based data has shed light on the biological roles of the epigenome [42–43]. There is considerable intra- and inter-species variation in DNA methylation patterns. The analysis of this type of data is not limited to model plant species with compact genomes, instead extending to important agronomic crops with large, complex genomes [34, 44–47]. Natural genomic variations, such as single-nucleotide mutations and structural variations, have been used as tools in plant breeding [34, 47].

In the current study, we determined that the average methylation levels of all *B. napus* samples (from high to low) were mCG (43.99%), mCHG (29.09%), and mCHH (26.92%). A recent study indicated that a *B. napus* male sterile line (7365A) and a restorer line (7365B) showed completely different methylation patterns (Fig. 14a and b) [21]. There was no significant difference in the methylation levels of CG, CHG, and CHH between the sterile line and the restorer line, but for both lines, the overall methylation level was mCHH > mCG > mCHG. This finding indicates that changes in methylation patterns are strongly related to the fertility of *B. napus*.

In the current study, all of the *B. napus* samples used for sequencing showed the same trend: the methylation levels of all three sequence contexts were significantly higher on chromosome C than on chromosome A. This result is consistent with previous studies [21, 48]. The A and C chromosomes of *B. napus* come from Chinese cabbage (*Brassica rapa*) and *Brassica oleracea*, respectively [48–50]. Based on separate studies, the overall methylation level of the *Brassica oleracea* genome is significantly higher than that of *Brassica rapa* [51–52]. Together, these findings further support the notion that *Brassica napus* evolved from the hybridization of *B. rapa* and *B. oleracea* and that the methylation patterns of parental species *B. rapa* and *B. oleracea* were stably inherited by their offspring, *B. napus*.

Through joint analysis of methylation and transcriptome data, we determined that the promoter regions of many important genes are hypermethylated or demethylated, which affects the differential expression of the genes between the high- and low-TPC materials. Due to hypermethylation in the promoter region of SDH2-2, the expression of this gene was reduced in both the stem apex and buds of the high-TPC lines. Indeed, increased methylation in the promoter region affects the expression of SDH2-2 in maize seeds, thereby affecting the glyoxylic acid cycle and seed germination [53]. Therefore, the differential methylation of the SDH2-2 promoter region causes significant differences in gene expression between the high- and low-TPC materials, thereby affecting the glyoxylic acid cycle, leading to differences in the characteristics of the two groups of materials.

NRT1.8 is hypermethylated in both the stem apex and buds of the high TPC lines, leading to significantly reduced gene expression. NRT1.8 affect nitrogen use efficiency in *Brassica napus*, and nitrogen has major effects on TPC-related traits [54–55]. The loss of methylation in the promoter regions of CYSB and UBC32 cause these genes to be expressed at high levels in the high-TPC lines. CYSB is involved in the nitrogen stress response in spinach [56]. The Arabidopsis ubiquitin-conjugating enzyme UBC32 is an ERAD element that plays an important role in brassinolide-mediated plant growth and salt stress tolerance [57]. In the flower buds of the high-TPC materials, the TAF15b promoter was hypomethylated, leading to significantly higher gene expression compared to the low-TPC lines. TAF15b participates in the autonomous pathway of flowers, inhibits the transcription of FLOWERING LOCUS C, and regulates the development of floral organs [58–59]. Floral organ development plays an important role in silique development, which in turn is important for determining the TPC trait.

RdDM is an important epigenetic pathway that affects many traits in plants [60–61]. MicroRNA-mediated methylation is widely involved in regulating gene function in both animals and plants. In humans, miR-29 targets DNA demethylase genes, and its activity is mediated by members of the ten eleven translocation (TET) family, which may cause the dysregulation of genes involved in key cell functions, leading to lung cancer [62]. MiR-140-5p regulates T cell differentiation and reduces autoimmune encephalomyelitis by affecting CD4 + T cell metabolism and DNA methylation [63]. miR2936 and miR398-mediated DNA methylation affect respiratory energy metabolism in Arabidopsis by regulating the expression of AGO1 and AGO4 [26]. Floral organ development plays a vital role in plant reproduction and the TPC trait. The RdDM pathway mediated by microRNAs regulates the poly(A) polymerase gene PAPS1 in Arabidopsis, thus affecting both sporophyte and pollen development [28]. Therefore, miRNA-mediated DNA methylation plays an important role in the normal growth and development of animals and plants.

Here, we compared genome-wide methylation, transcriptome, and miRNA data and found that many gene promoter regions in the high- and low-TPC materials were jointly regulated by miRNAs and methylation, leading to significant differential gene expression. These genes are closely related to the TPC trait. For example, the BnaCnng64040D (Lipase family protein) promoter is regulated by miR159b and is hypermethylated in flower buds of the high-TPC materials, leading to significantly reduced expression. Lipase family protein plays an important role in regulating lipid metabolism and oil content of *B. napus* [64]. The promoter region of BnaC09g30490D (TAF15b), which affects organ development, is regulated by miR167c and undergoes demethylation in the high-TPC materials, leading to significantly higher expression in flower buds compared to the low-TPC materials. The promoter region of BnaC04g55380D (MAP65-4) is regulated by miR6029/miR9049 and is hypermethylated, which may affect the formation of the cytoskeleton in *B. napus* [65]. In fact, we found that together with differences in methylation, many miRNAs that are differentially expressed between high- and low-TPC lines are closely related to the formation of TPC, such as miR159, miR827, miR319, and so on [1].

Finally, GO and KEGG analysis of genes that are regulated by methylation in their promoter regions and are differentially expressed in high- vs. low-TPC materials primarily function in metabolic processes such as carbon and nitrogen metabolism. We previously reported that miRNAs are involved in the differential expression of genes between high- and low-TPC lines. GO and KEGG analysis revealed that these genes are also mainly involved in carbon and nitrogen metabolism [1]. Therefore, the results obtained in the current study are consistent with previous results. These findings indicate that the TPC trait is simultaneously regulated by miRNAs and methylation and is closely related to carbon and nitrogen metabolism processes.

DNA methylation is an epigenetic modification that affects gene expression and transposable element activity. Because cytosine DNA methylation patterns are inherited through mitotic and meiotic cell division, differences in these patterns may lead to phenotypic variation. Advances in high-throughput sequencing technology have led to the generation of abundant DNA sequence data. The comprehensive analysis of genome-wide gene expression and DNA methylation patterns has revealed the underlying mechanisms and functions of DNA methylation. In addition, various associations between DNA methylation and agronomic traits have been uncovered [66]. The information obtained from such studies could be used for crop breeding based on natural epigenomic variations in the future. In addition, artificial epigenome editing could be used as a new breeding technique to produce new crop varieties with improved agronomic traits.

5. Conclusions

Significant differences were detected in overall methylation levels between the high- and low-TPC lines in the CG, CHG, and CHH contexts in the promoters of genes in the stem apex and flower bud. In addition, 14 important genes related to growth and development were differentially regulated between the two groups due to miRNA-mediated differences in methylation levels in their promoters. These results, together with our previously generated RNA-seq and miRNA profiling data, indicate that both methylation and miRNAs are involved in regulating the expression of genes in nitrogen-related metabolic pathways, thereby affecting the TPC trait in *B. napus*.

Declarations

Acknowledgments

We sincerely thank Prof. Jiana Li for laboratory assistance and helpful discussions.

Authors' contributions

This work was contributed by Z. C., L. J. and Y. W. equally. Z.C. and J.L. designed the research study; Z.C., L.J., C.Q., and K.L. performed the bioinformatics analysis; Y.W. and J.M. performed the validation experiments and analysis. All authors have read and approved of the final version of the manuscript.

Funding

This work was supported by the National Key Research and Development Plan (2018YFD0100504), the National Natural Science Foundation of China (31830067 and U1302266), the 973 Project (2015CB150201), and the 111 Project (B12006).

Availability of data and materials

The datasets during the current study are deposited in the publicly accessible NCBI Sequence Read Archive (SRA) Database as accession numbers SRR8477015, SRR8477016, SRR8477023, SRR8477024, SRR8477025, SRR8477026, SRR8477027, and SRR8477028.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Z. Chen, Q. Huo, H. Yang, H. Jian, C. Qu, K. Lu, J. Li, Joint RNA-Seq and miRNA Profiling Analyses to Reveal Molecular Mechanisms in Regulating Thickness of Pod Canopy in *Brassica napus*, *Genes* 10 (2019) 591.
2. Y. Zhang, C. Yu, J. Lin, J. Liu, B. Liu, J. Wang, A. Huang, H. Li, T. Zhao, OsMPH1 regulates plant height and improves grain yield in rice, *PLoS One* 12 (2017) e0180825.
3. W.T. Liu, P.W. Chen, L.C. Chen, C.C. Yang, S.Y. Chen, G. Huang, T. Lin, H.M. Ku, J. Chen, Suppressive effect of microRNA319 expression on rice plant height, *Theor. Appl. Genet.* 130 (2017) 1507–1518.
4. F. Li, B. Chen, K. Xu, G. Gao, G. Yan, J. Qiao, J. Li, H. Li, L. Li, X. Xiao, T. Zhang, T. Nishio, X. Wu, A genome-wide association study of plant height and primary branch number in rapeseed (*Brassica napus*), *Plant Science* 242 (2016) 169–177.
5. M. Zheng, C. Peng, H. Liu, M. Tang, H. Yang, X. Li, J. Liu, X. Sun, X. Wang, J. Xu, W. Hua, H. Wang, Genome-Wide Association Study Reveals Candidate Genes for Control of Plant Height, Branch Initiation Height and Branch Number in Rapeseed (*Brassica napus* L.), *Front. Plant Sci.* 8 (2017) 1246.
6. P. Satyaki, M. Gehring, DNA methylation and imprinting in plants: machinery and mechanisms, *Crit. Rev. Biochem. Mol. Biol.* 52 (2017) 163–175.
7. Y. Gruenbaum, T. Naveh-Many, H. Cedar, A. Razin, Sequence specificity of methylation in higher plant DNA. *Nature* 292 (1981) 860–862.
8. P. Meyer, I. Niedenhof, M. Lohuis, Evidence for cytosine methylation of non-symmetrical sequences in transgenic *Petunia hybrida*, *EMBO J.* 13 (1994) 2084–2088.
9. J.A. Law, S.E. Jacobsen, Establishing, maintaining and modifying DNA methylation patterns in plants and animals, *Nat. Rev. Genet.* 11 (2010) 204–220.
10. M. Gehring, J. Huh, T.F. Hsieh, J. Penterman, Y. Choi, J. Harada, R. Goldberg, R. Fischer, DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation, *Cell* 124 (2006) 495–506.
11. A. Zemach, M.Y. Kim, P. Silva, J.A. Rodrigues, B. Dotson, M.D. Brooks, D. Zilberman, Local DNA hypomethylation activates genes in rice endosperm, *Proc. Natl. Acad. Sci. USA.* 107 (2010) 18729–18734.
12. Z. Lang, Y. Wang, K. Tang, D. Tang, T. Datsenka, J. Cheng, Y. Zhang, A.K. Handa, J.K. Zhu, Critical roles of DNA demethylation in the activation of ripening-induced genes and inhibition of ripening-repressed genes in tomato fruit, *Proc. Natl. Acad. Sci. USA.* 114 (2017) E4511-E4519.
13. H. Huang, R. Liu, Q. Niu, K. Tang, B. Zhang, H. Zhang, K. Chen, J.K. Zhu, Z. Lang, Global increase in DNA methylation during orange fruit development and ripening, *Proc. Natl. Acad. Sci. USA.* 116 (2019) 1430–1436.
14. W. Wang, Q. Qin, F. Sun, Y. Wang, D. Xu, Z. Li, B. Fu, Genome-Wide Differences in DNA Methylation Changes in Two Contrasting Rice Genotypes in Response to Drought Conditions, *Front. Plant Sci.* 7 (2016) 1675.
15. S.-J. Feng, X.-S. Liu, H. Tao, S.-K. Tan, S.-S. Chu, Y. Oono, X.-D. Zhang, J. Chen, Z.-M. Yang, Variation of DNA methylation patterns associated with gene expression in rice (*Oryza sativa*) exposed to cadmium, *Plant Cell Environ.* 39 (2016) 2629–2649.
16. L. Ji, S. Mathioni, S. Johnson, D. Tucker, A.-J. Bewick, K.-D. Kim, J. Daron, R.K. Slotkin, S.A. Jackson, W.A. Parrott, B.C. Meyers, R. Schmitz, Genome-wide reinforcement of DNA methylation occurs during somatic embryogenesis in soybean, *Plant Cell* 31 (2019) 2315–2331.
17. O.P. Gupta, A. Dahuja, A. Sachdev, P.K. Jain, T.V. Kumari, S. Praveen, Cytosine Methylation of Isoflavone Synthase Gene in the Genic Region Positively Regulates Its Expression and Isoflavone Biosynthesis in Soybean Seeds, *DNA Cell Biol.* 38 (2019) 510–520.
18. S. Mager, B. Schönberger, U. Ludwig, The transcriptome of zinc deficient maize roots and its relationship to DNA methylation loss, *BMC Plant Biol.* 18 (2018) 372.
19. Y. Qian, W. Hu, J. Liao, J. Zhang, Q. Ren, The Dynamics of DNA methylation in the maize (*Zea mays* L.) inbred line B73 response to heat stress at the seedling stage, *Biochem. Biophys. Res. Commun.* 512 (2019) 742–749.
20. S. Shafiq, Q. Zeb, A. Ali, Y. Sajjad, R. Nazir, E. Widemann, L. Liu, Lead, Cadmium and Zinc Phytotoxicity Alter DNA Methylation Levels to Confer Heavy Metal Tolerance in Wheat, *Int. J. Mol. Sci.* 20 (2019) 4676.
21. Z. Wang, X. Wu, Z. Wu, H. An, B. Yi, J. Wen, C. Ma, J. Shen, T. Fu, J. Tu, Genome-Wide DNA Methylation Comparison between *Brassica napus* Genic Male Sterile Line and Restorer Line, *Int. J. Mol. Sci.* 19 (2018) 2689.
22. J. Li, Q. Huang, M. Sun, T. Zhang, H. Li, B. Chen, K. Xu, G. Gao, F. Li, G. Yan, J. Qiao, Y. Cai, X. Wu, Global DNA methylation variations after short-term heat shock treatment in cultured microspores of *Brassica napus* cv. Topas, *Sci. Rep.* 6 (2016) 38401.

23. P. Jiang, S. Wang, H. Jiang, B. Cheng, K. Wu, Y. Ding, The COMPASS-Like Complex Promotes Flowering and Panicle Branching in Rice, *Plant Physiol.* 176 (2018) 2761–2771.
24. X. Liu, R. Wei, Z. Sheng, G. Jiao, S. Tang, J. Luo, P. Hu, Polycomb Protein OsFIE2 Affects Plant Height and Grain Yield in Rice, *PLoS ONE* 11 (2016) e0164748.
25. Y. Hu, G. Morota, G. Rosa, D. Gianola, Prediction of Plant Height in *Arabidopsis thaliana* Using DNA Methylation Data, *Genetics* 201 (2015) 779–793.
26. E. Loreti, F. Betti, C.M. Ladera, F. Fontana, G. Novi, M.C. Valeri, P. Perata, ARGONAUTE1 and ARGONAUTE4 regulate gene expression and hypoxia tolerance, *Plant Physiol.* 182 (2020) 287–300.
27. B. Wang, J. Liu, L. Chu, X. Jing, H. Wang, J. Guo, B. Yi, Exogenous Promoter Triggers APETALA3 Silencing through RNA-Directed DNA Methylation Pathway in *Arabidopsis*, *Int. J. Mol. Sci.* 20 (2019) 4478.
28. Y. Zhang, A. Ramming, L. Heinke, L. Altschmied, R.K. Slotkin, J.D. Becker, C. Kappel, M. Lenhard, The poly(A) polymerase PAPS1 interacts with the RNA-directed DNA-methylation pathway in sporophyte and pollen development, *Plant J.* 99 (2019) 655–672.
29. M. Iwasaki, L. Hyvärinen, U. Piskurewicz, M.L. Lopez, Non-canonical RNA-directed DNA methylation participates in maternal and environmental control of seed dormancy, *ELife* 8 (2019) e37434.
30. X. Zhong, Z.Q. Wang, R. Xiao, Y. Wang, Y. Xie, X. Zhou, iTRAQ analysis of the tobacco leaf proteome reveals that RNA-directed DNA methylation (RdDM) has important roles in defense against geminivirus-betasatellite infection, *J. Proteomics.* 152 (2017) 88–101.
31. E. Gruntman, Y. Qi, R.K. Slotkin, T. Roeder, R.A. Martienssen, R. Sachidanandam, Kismeth: Analyzer of plant methylation states through bisulfite sequencing, *BMC Bioinformatics* 9 (2008) 371.
32. J. Wang, C. Wang, Y. Long, C. Hopkins, S. Kurup, K. Liu, G.J. King, J. Meng, Universal endogenous gene controls for bisulphite conversion in analysis of plant DNA methylation, *Plant Methods* 7 (2011) 39.
33. M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D.J. Horsman, S. Jones, M. Marra, CIRCOS: an information aesthetic for comparative genomics, *Genome Res.* 19 (2009) 1639–1645.
34. S. Zhong, Z. Fei, Y.R. Chen, Y. Zheng, M. Huang, J. Vrebalov, R. McQuinn, N. Gapper, B. Liu, J. Xiang, Y. Shao, J.J. Giovannoni, Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening, *Nat. Biotechnol.* 31 (2013) 154–159.
35. M. Kulis, S. Heath, M. Bibikova, A.C. Queirós, A. Navarro, G. Clot, A. Martínez-Trillos, G. Castellano, I. Brun-Heath, M. Pinyol, S. Barberán-Soler, P. Papasaikas, P. Jares, S. Beà, D. Rico, S. Ecker, M. Rubio, R. Royo, V. Ho, B. Klotzle, L. Hernández, L. Conde, M. López-Guerra, D. Colomer, N. Villamor, M. Aymerich, M. Rozman, M. Bayes, M. Gut, J.L. Gelpí, M. Orozco, J.B. Fan, V. Quesada, X.S. Puente, D.G. Pisano, A. Valencia, A. López-Guillermo, I. Gut, C. López-Otín, E. Campo, J.I. Martín-Subero, Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia, *Nat. Genet.* 44 (2012) 1236–1242.
36. H. Feng, K. Conneely, H. Wu, A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data, *Nucleic Acids Res.* 42 (2014) e69.
37. H. Wu, T. Xu, H. Feng, L. Chen, B. Li, B. Yao, Z. Qin, P. Jin, K. Conneely, Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates, *Nucleic Acids Res.* 43 (2015) e141.
38. Y. Park, H. Wu, Differential methylation analysis for BS-seq data under general experimental design, *Bioinformatics (Oxford, England)* 32 (2016) 1446–1453.
39. A. Bartels, Q. Han, P. Nair, L. Stacey, H. Gaynier, M. Mosley, Q. Huang, J. Pearson, T.F. Hsieh, Y.Q. An, W. Xiao, Dynamic DNA Methylation in Plant Growth and Development, *Int. J. Mol. Sci.* 19 (2018) 2144.
40. S.A. Smallwood, H.J. Lee, C. Angermueller, F. Krueger, H. Saadeh, J. Peat, S.R. Andrews, O. Stegle, W. Reik, G. Kelsey, Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity, *Nat. Methods.* 11 (2014) 817–820.
41. R. Lister, R.C. O'Malley, J. Tonti-Filippini, B.D. Gregory, C.C. Berry, A.H. Millar, J.R. Ecker, Highly Integrated Single-Base Resolution Maps of the Epigenome in *Arabidopsis*, *Cell* 133 (2008) 523–536.
42. T. Kawakatsu, S.S. Huang, F. Jupe, E. Sasaki, R. Schmitz, M. Urich, R. Castanon, J. Nery, C. Barraga, Y. He, H. Chen, M. Dubin, C.R. Lee, C. Wang, F. Bemm, C. Becker, R. Oneil, R.C. O'Malley, D.X. Quarless, B. Thanh, Epigenomic Diversity in a Global Collection of *Arabidopsis thaliana* Accessions, *Cell* 166 (2016) 492–505.
43. C. Niederhuth, A. Bewick, L. Ji, M. Alabady, K.D. Kim, Q. Li, N. Rohr, A. Rambani, J. Burke, J. Udall, C. Egesi, J. Schmutz, J. Grimwood, S. Jackson, N. Springer, R. Schmitz, Widespread natural variation of DNA methylation within angiosperms, *Genome Biol.* 17 (2016) 194.
44. M. Regulski, L. Zhenyuan, J. Kendall, M. Donoghue, J. Reinders, V. Llaca, S. Deschamps, A. Smith, D. Levy, W. McCombie, S. Tingey, A. Rafalski, J. Hicks, R. Martienssen, The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA, *Genome Res.* 23 (2013) 1651–1662.
45. R. Schmitz, Y. He, O. Valdés-López, S. Khan, T. Joshi, M. Urich, J. Nery, B. Diers, D. Xu, G. Stacey, J. Ecker, Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population, *Genome Res.* 23 (2013) 1663–1674.

46. N. Daccord, J.M. Celton, G. Linsmith, C. Becker, N. Choisine, E. Schijlen, H. van de Geest, L. Bianco, D. Micheletti, R. Velasco, E.A. Di Pierro, J. Gouzy, D.J.G. Rees, P. Guérif, H. Muranty, C.E. Durel, F. Laurens, Y. Lespinasse, S. Gaillard, S. Aubourg, H. Quesneville, D. Weigel, E. van de Weg, M. Troggo, E. Bucher, High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development, *Nat. Genet.* 49 (2017) 1099–1106.
47. G.M. Turco, K. Kajala, G. Kunde-Ramamoorthy, C.Y. Ngan, A. Olson, S. Deshpande, D. Tolkunov, B. Waring, S. Stelpflug, P. Klein, J. Schmutz, S. Kaeppler, D. Ware, C.L. Wei, J.P. Etchells, S.M. Brady, DNA methylation and gene expression regulation associated with vascularization in *Sorghum bicolor*, *New Phytol.* 214 (2017) 1213–1229.
48. B. Chalhoub, F. Denoeud, S. Liu, I.A.P. Parkin, H. Tang, X. Wang, J. Chiquet, H. Belcram, C. Tong, B. Samans, M. Corr ea, C. Da Silva, J. Just, C. Falentin, C.S. Koh, I. Le Clainche, M. Bernard, P. Bento, B. Noel, K. Labadie, A. Alberti, M. Charles, D. Arnaud, H. Guo, C. Daviaud, S. Alamery, K. Jabbari, M. Zhao, P.P. Edger, H. Chelaifa, D. Tack, G. Lassalle, I. Mestiri, N. Schnell, M.C. Le Paslier, G. Fan, V. Renault, P.E. Bayer, A.A. Golicz, S. Manoli, T.H. Lee, V.H.D. Thi, S. Chalabi, Q. Hu, C. Fan, R. Tollenaere, Y. Lu, C. Battail, J. Shen, C.H.D. Sidebottom, X. Wang, A. Canaguier, A. Chauveau, A. B erard, G. Deniot, M. Guan, Z. Liu, F. Sun, Y.P. Lim, E. Lyons, C.D. Town, I. Bancroft, X. Wang, J. Meng, J. Ma, J.C. Pires, G.J. King, D. Brunel, R. Delourme, M. Renard, J.M. Aury, K.L. Adams, J. Batley, R.J. Snowdon, J. Tost, D. Edwards, Y. Zhou, W. Hua, A.G. Sharpe, A.H. Paterson, C. Guan, P. Wincker, Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome, *Science* 345 (2014) 950–953.
49. U. Nagaharu, Genome analysis in *Brassica* with special reference to the experimental formation of *B. Napus* and peculiar mode of fertilization, *Japan Journal of Botany* 7 (1935) 389–452.
50. C. Allender, G. King, Origins of the amphiploid species *Brassica napus* L. investigated by chloroplast and nuclear molecular markers, *BMC Plant Biol.* 10 (2010) 54.
51. I.A.P. Parkin, C. Koh, H. Tang, S.J. Robinson, S. Kagale, W.E. Clarke, C.D. Town, J. Nixon, V. Krishnakumar, S.L. Bidwell, F. Denoeud, H. Belcram, M.G. Links, J. Just, C. Clarke, T. Bender, T. Huebert, A.S. Mason, J.C. Pires, G. Barker, J. Moore, P.G. Walley, S. Manoli, J. Batley, D. Edwards, M.N. Nelson, X. Wang, A.H. Paterson, G. King, I. Bancroft, B. Chalhoub, A.G. Sharpe, Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*, *Genome Biol.* 15 (2014) R77.
52. Y. Liu, C. Xu, X. Tang, S. Pei, D. Jin, M. Guo, M. Yang, Y. Zhang, Genomic methylation and transcriptomic profiling provides insights into heading depression in inbred *Brassica rapa* L. ssp. *Pekinensis*, *Gene* 665 (2018) 119–126.
53. A. Eprintsev, D. Fedorin, L. Karabutova, A. Igamberdiev, Expression of genes encoding subunits A and B of succinate dehydrogenase in germinating maize seeds is regulated by methylation of their promoters, *J. Plant Physiol.* 205 (2016) 33–40.
54. Y.L. Han, H.X. Song, Q. Liao, Y. Yu, S.F. Jian, J.E. Lepo, Q. Liu, X.M. Rong, C. Tian, J. Zeng, C.Y. Guan, A.M. Ismail, Z.H. Zhang, Nitrogen Use Efficiency Is Mediated by Vacuolar Nitrate Sequestration Capacity in Roots of *Brassica napus*, *Plant Physiol.* 170 (2016) 1684–1698.
55. S. Khan, S. Anwar, J. Kuai, S. Ullah, S. Fahad, G. Zhou, Optimization of Nitrogen Rate and Planting Density for Improving Yield, Nitrogen Use Efficiency, and Lodging Resistance in Oilseed Rape, *Front. Plant. Sci.* 8 (2017) 532–532.
56. H. Takahashi, K. Saito, Subcellular localization of spinach cysteine synthase isoforms and regulation of their gene expression by nitrogen and sulfur, *Plant Physiol.* 112 (1996) 273–280.
57. F. Cui, L. Liu, Q. Zhao, Z. Zhang, Q. Li, B. Lin, Y. Wu, S. Tang, Q. Xie, Arabidopsis Ubiquitin Conjugase UBC32 Is an ERAD Component That Functions in Brassinosteroid-Mediated Salt Stress Tolerance, *Plant Cell.* 24 (2012) 233–244.
58. H. Eom, S. Park, M. Kim, H. Kim, H. Kang, I. Lee, TAF15b, involved in the autonomous pathway for flowering, represses transcription of FLOWERING LOCUS C, *Plant J.* 93 (2017) 79–91.
59. H. Eom, I. Lee, Role of TAF15b in transcriptional regulation of autonomous pathway for flowering, *Plant. Signal. Behav.* 13 (2018) 1–3.
60. M. Matzke, R. Mosher, RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* 15 (2014) 394–408.
61. D. Cuerda-Gil, R. Slotkin, Non-canonical RNA-directed DNA methylation, *Nat. Plants.* 2 (2016) 16163.
62. B. Humphries, Z. Wang, C. Yang, MicroRNA Regulation of Epigenetic Modifiers in Breast Cancer, *Cancers* 11 (2019) 897.
63. S. Zhu, X. Zhang, H. Guan, F. Huang, L. Wu, D. Hou, Z. Zheng, M. Yu, L. Huang, L. Ge, miR-140-5p regulates T cell differentiation and attenuates experimental autoimmune encephalomyelitis by affecting CD4⁺T cell metabolism and DNA methylation, *Int. Immunopharmacol.* 75 (2019) 105778.
64. A. Kelly, E. Shaw, S.J. Powers, S. Kurup, P. Eastmond, Suppression of the SUGAR-DEPENDENT1 triacylglycerol lipase family during seed development enhances oil yield in oilseed rape (*Brassica napus* L.), *Plant Biotechnol J.* 11 (2012) 355–361.
65. H. Li, B. Sun, M. Sasabe, X. Deng, Y. Machida, H. Lin, Y.R. Julie Lee, B. Liu, Arabidopsis MAP65-4 plays a role in phragmoplast microtubule organization and marks the cortical cell division site, *New Phytol.* 215 (2017) 187–201.
66. T. Kawakatsu, J.Ecker, Diversity and dynamics of DNA methylation: epigenomic resources and tools for crop breeding, *Breed. Sci.* 69 (2019) 191–204.

Figures

YC4F_1

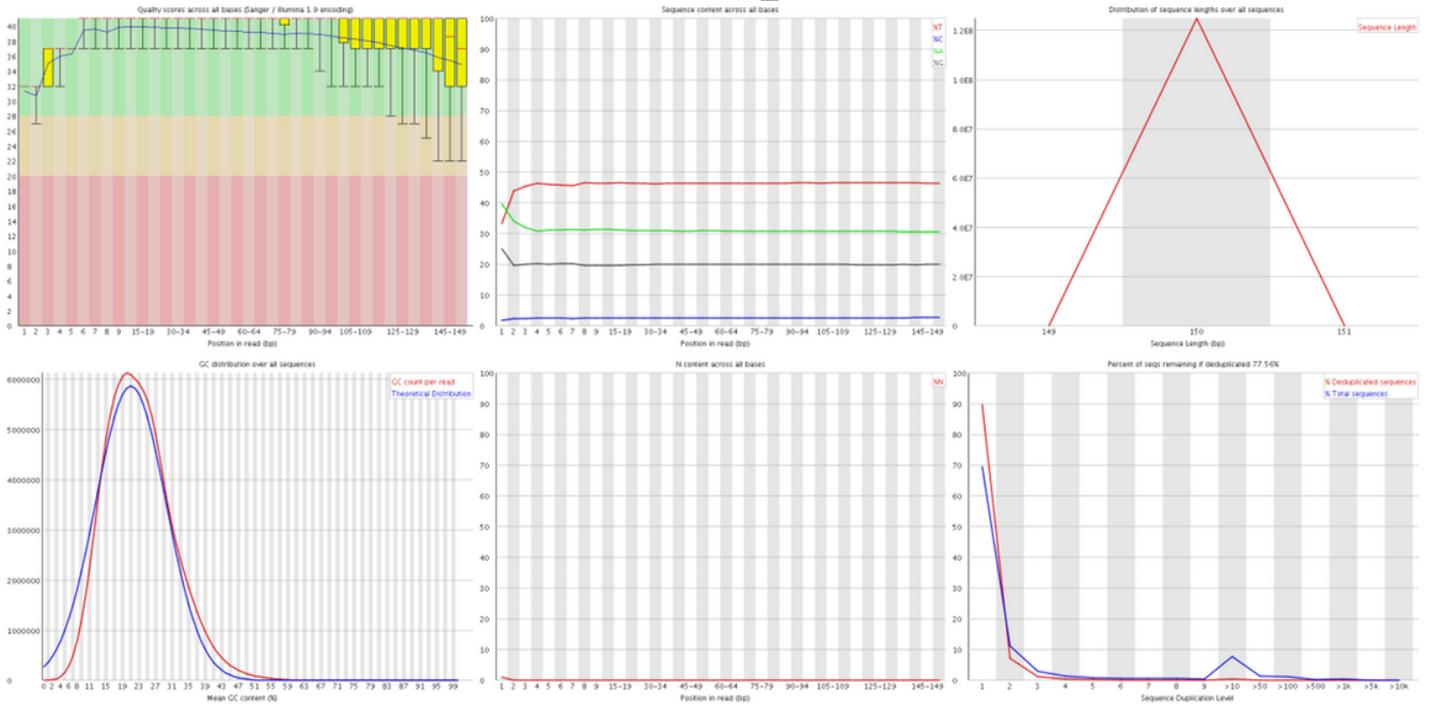


Figure 1

Raw reads of FastQC partial samples. Top left: The horizontal axis represents the position, the vertical axis quality, and the color distinguishes different mass segments; the upper middle image: the horizontal axis is the position, the vertical axis is the base percentage; the upper right picture: the distribution of the read length; the lower left picture: The distribution of the average GC content of the statistics reads, the red line is the actual situation, the blue line is the theoretical distribution; the lower middle picture: the ratio of the statistical N for each position of all reads; the lower right picture: the number of times the abscissa is duplication, the ordinate is the number of duplicated reads.

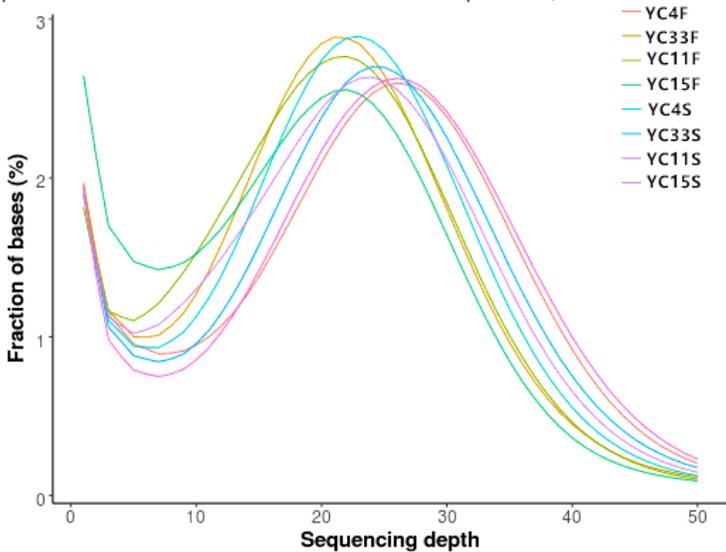


Figure 2

Genomic coverage map of all samples. Above: Distribution map, the abscissa is the depth of coverage, and the ordinate is the ratio of the number of base sites corresponding to the coverage to the number of bases in the whole genome; the following figure: the cumulative distribution map, the abscissa is the coverage depth, and the ordinate is greater than The ratio of the number of base sites equal to the corresponding coverage to the number of bases in the entire genome; the lines of different colors represent different samples.

YC4F

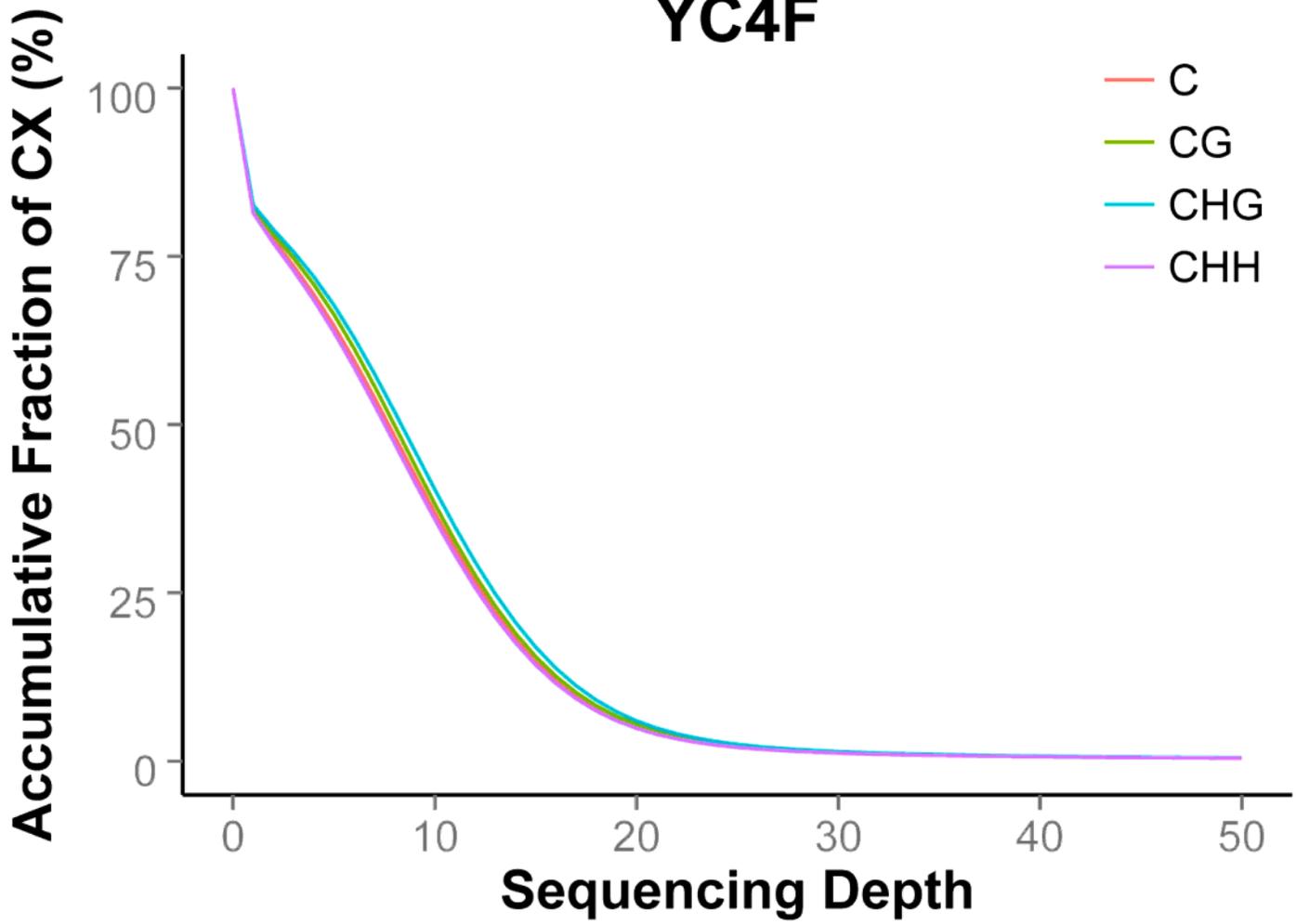


Figure 3

Partial cumulative distribution map of C-site coverage. The abscissa is the coverage depth of the C site, and the ordinate is the percentage of the C site that is greater than or equal to the corresponding coverage depth as a percentage of the total C site, and the different colors represent different contexts.

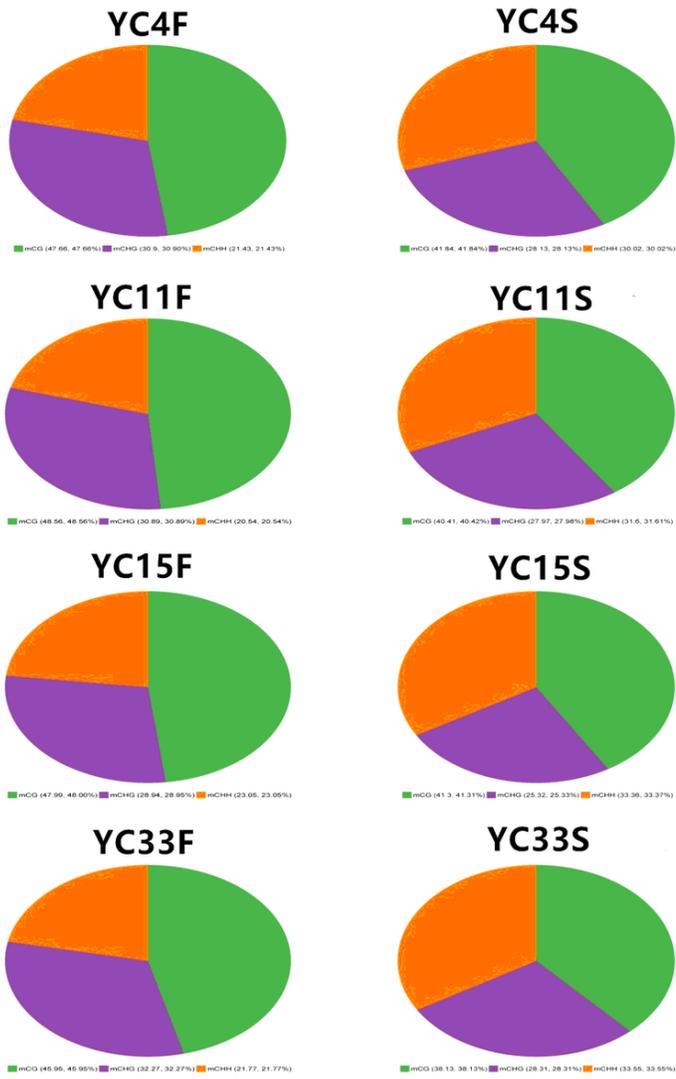


Figure 4

Proportional distribution map of methylation C site. Different colors represent methylation C sites under different contexts, and the size of each part represents the proportion of methylated C sites under the corresponding context.

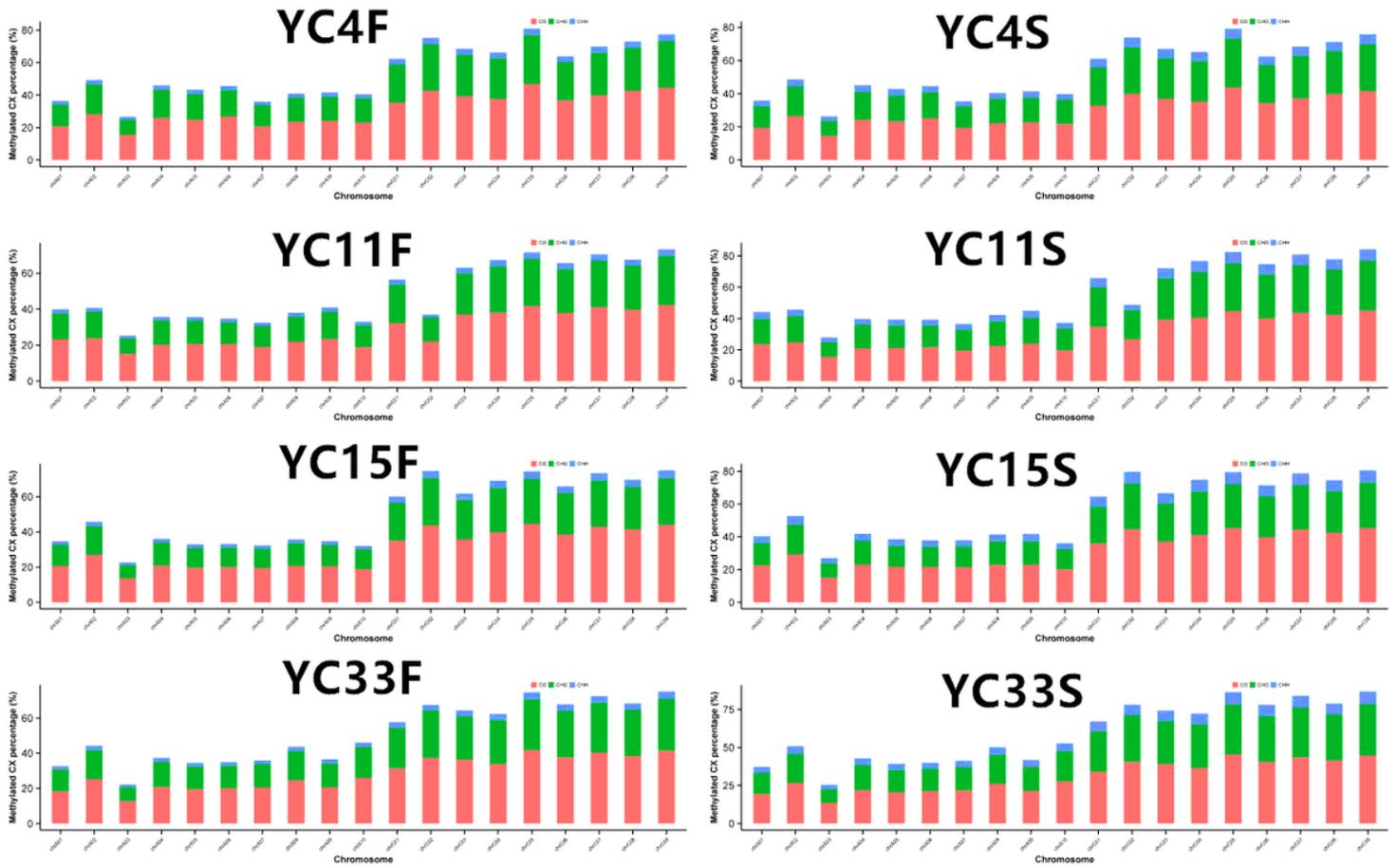


Figure 5

Percentage of C-sites methylation for each sequence environment at the chromosome level. Different colors represent methylation C sites under different contexts, and the length of each column represents the percentage of the sequence methylation site of the sequence to the sequence environment of the chromosome.

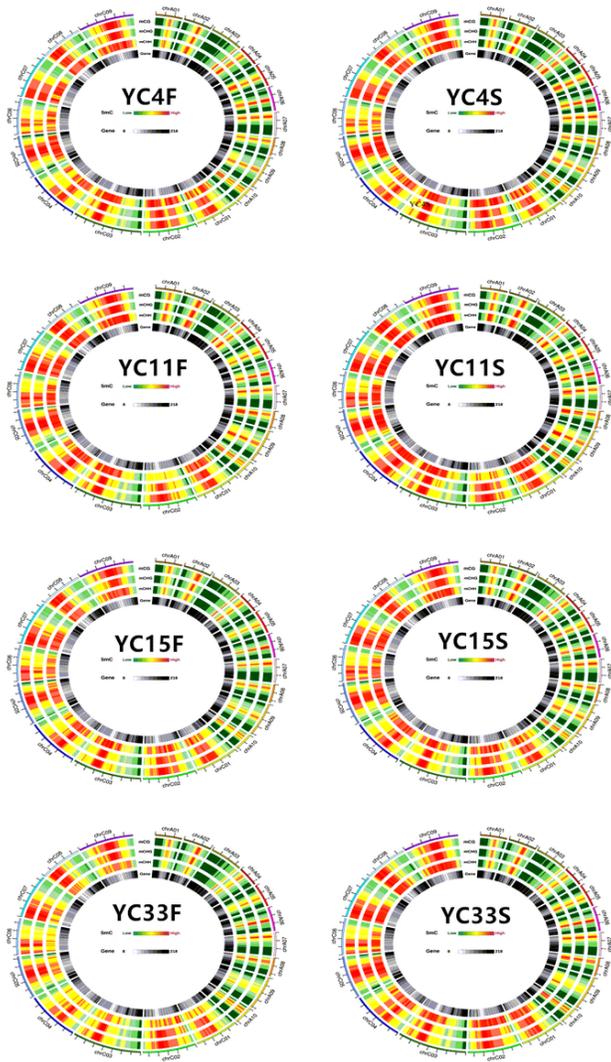


Figure 6

Circos diagram of chromosome methylation density. From the outside to the inside, the CG sequence environment methylation density, CHG sequence environment methylation density, CHH sequence environment methylation density, TE original proportion density heat map, gene number density heat map; internal scale: MC density Hot label: from green to yellow to red, the methylation density is from low to high, and the TE ratio is hot: from green to black to red, the repeat sequence ratio is from low to high, and the gene density is hot: from gray to black. Indicates the number of genes from low to high. TE density calculation method: Calculate the TE (repeat original) ratio, which is the total length ratio of the repeat in the bin (the analysis is not performed without repeat); Gene density calculation: Calculate the number of genes contained in each bin.

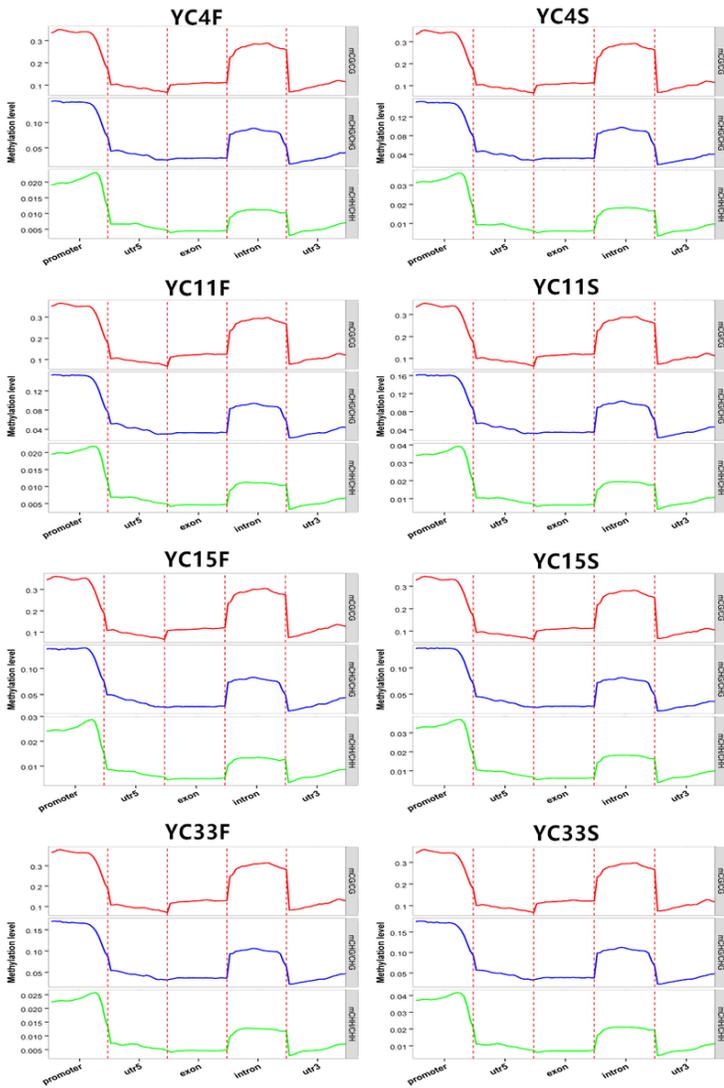


Figure 7
 Distribution of sample methylation levels on different genomic elements. The abscissa represents different genomic elements and the ordinate represents the level of methylation. The functional regions of each gene were equally divided into 20 bins, and then the C-site levels of the corresponding bins of the functional regions of all genes were averaged, and the different colors represented different sequence contexts (CpG, CHG, CHH).

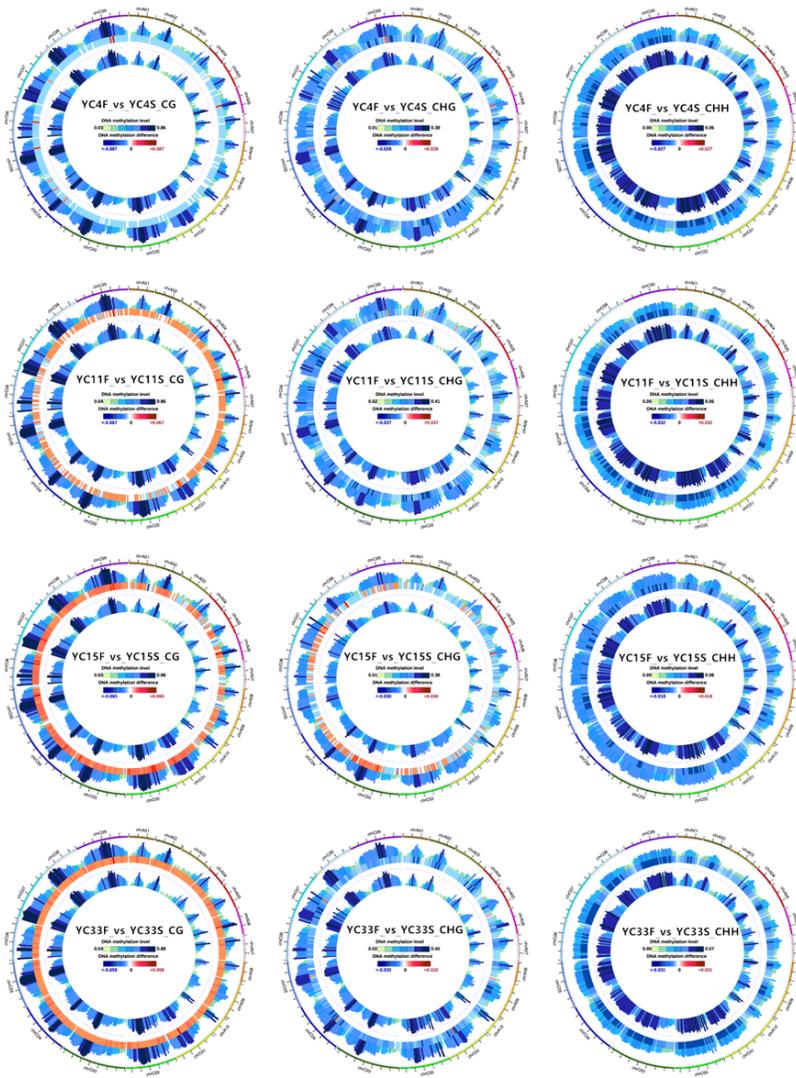


Figure 8

Difference in methylation level between stem and flower bud. From the outside to the inside, the circle indicates the methylation level of the treatment group, the difference of methylation level between the sample groups, and the methylation level of the control group. The internal scale: the DNA methylation level indicates the level of methylation. The DNA methylation difference indicates the degree of difference in methylation levels between samples.

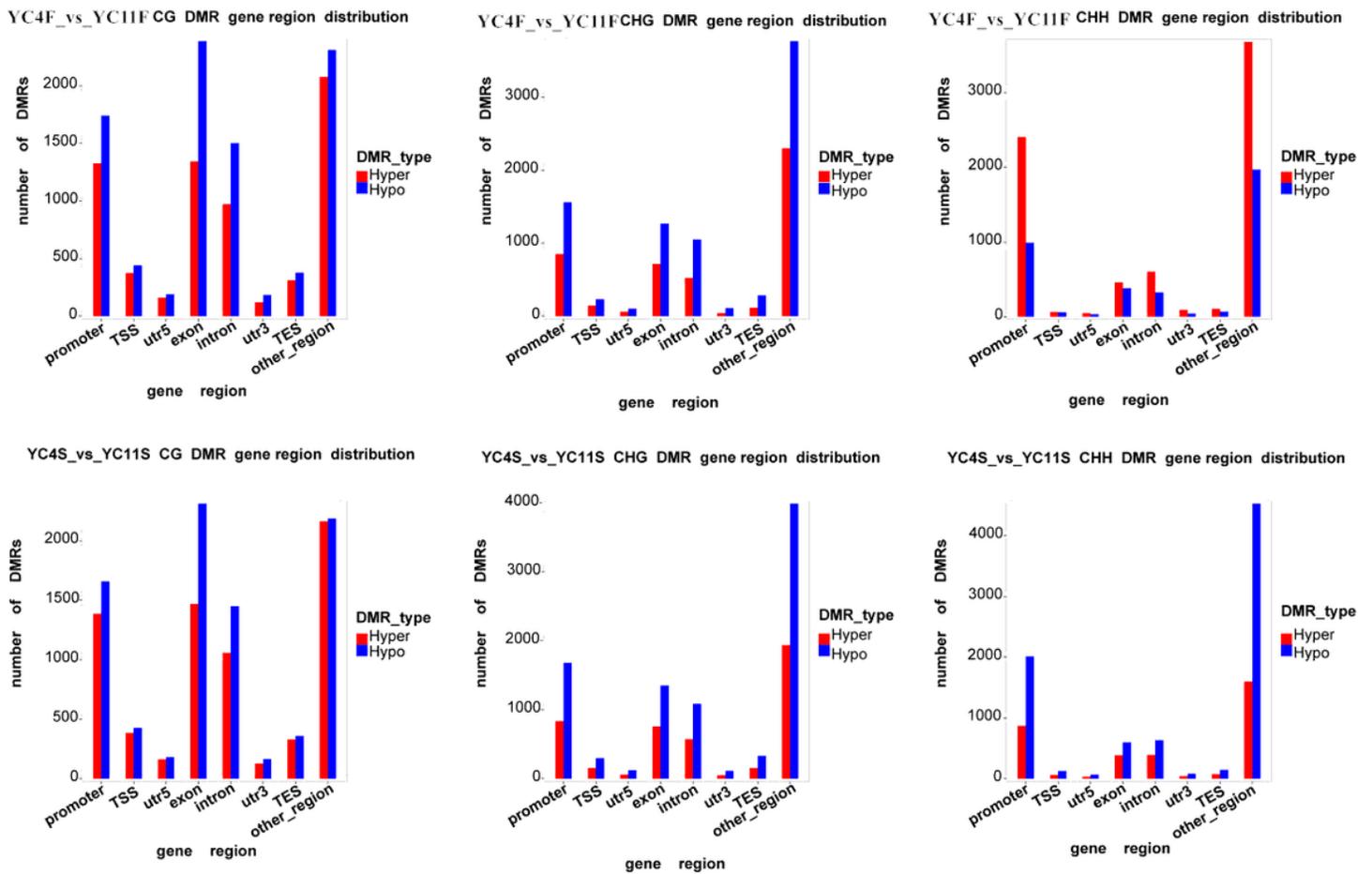


Figure 9

DMR anchoring area display of three sequence environment (CG, CHG, CHH). The abscissa represents the respective area categories, and the ordinate represents the number of DMRs of the hyper/hypo DMR in each area.

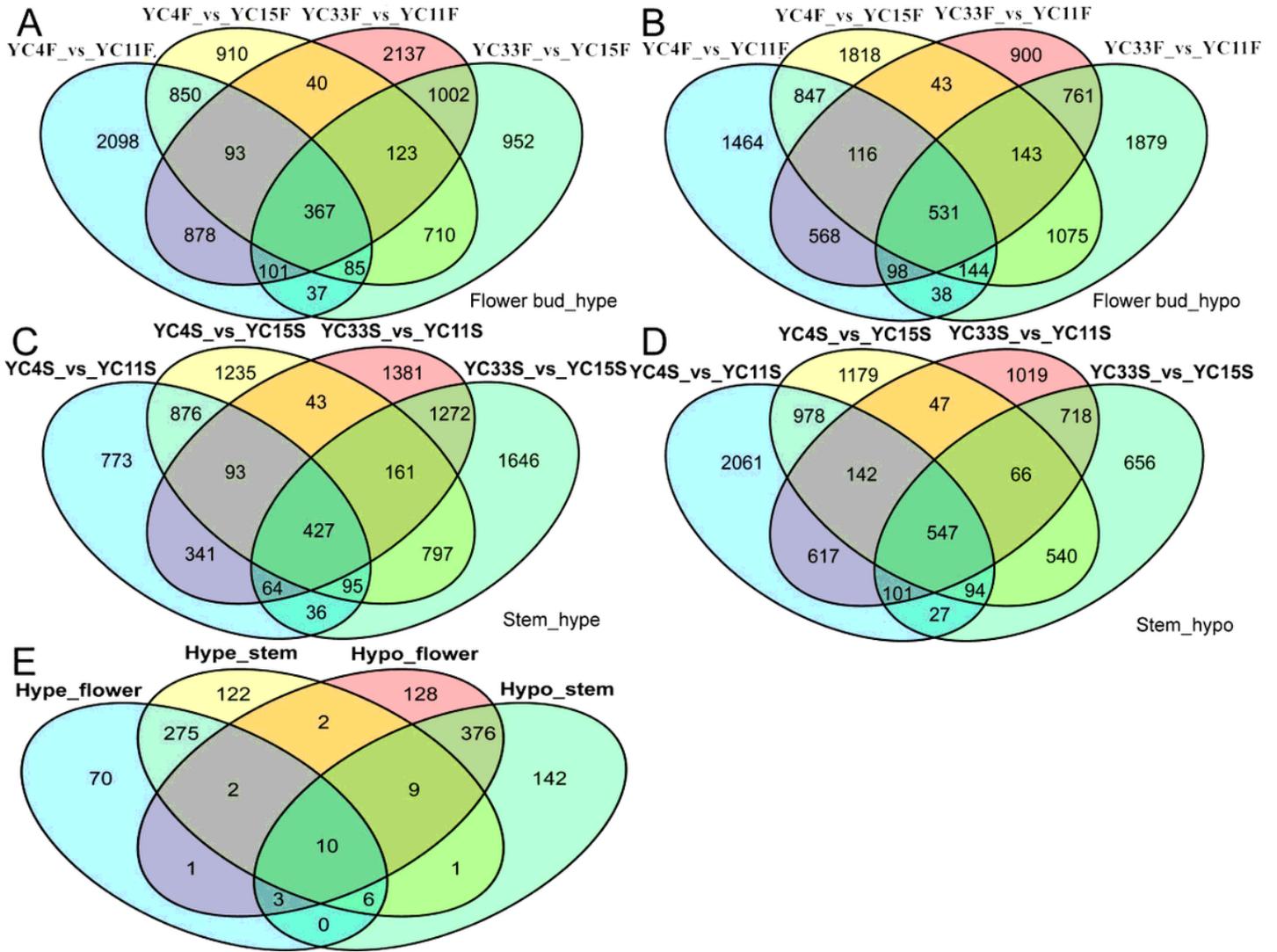


Figure 10

Differential gene for promoter methylation between high- and low-TPC materials. A: Genes with hypermethylation at flower bud in high TPC materials compared to low TPC materials; B: Genes with hypomethylation at flower bud in high TPC materials compared to low TPC materials; C: Genes with hypermethylation at stem apex in high TPC materials compared to low TPC materials; D: Genes with hypomethylation at stem apex in high TPC materials compared to low TPC materials. E: All high- and low-methylated genes in the promoter region at both sites of two groups of materials.

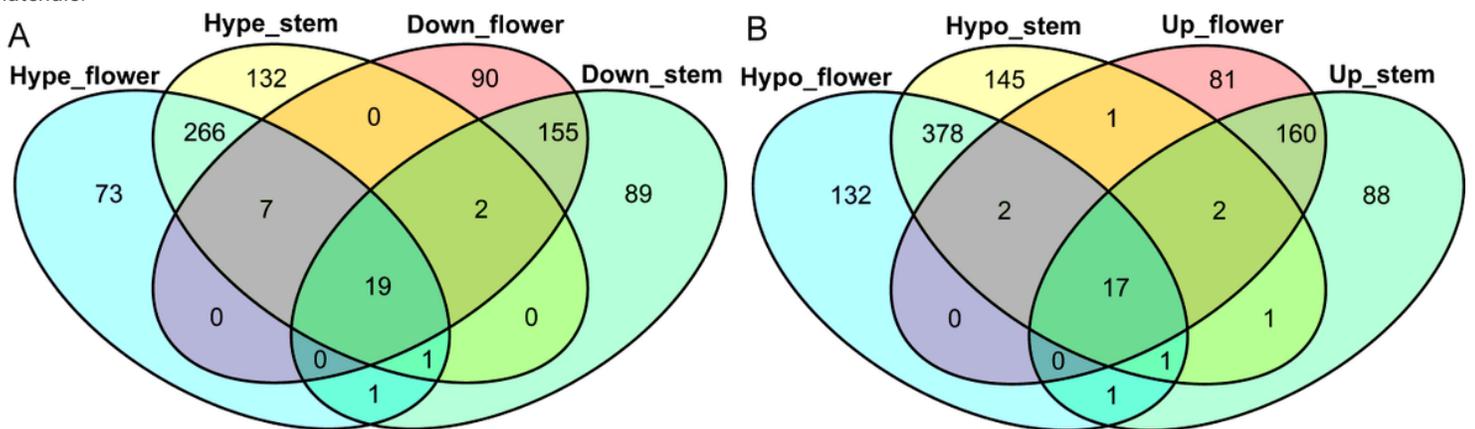


Figure 11

Methylation affects gene expression between high- and low- TPC materials in *Brassica napus* L. A: Hypermethylation in the promoter region results in a significant decrease in gene expression in the high-TPC materials compared to the low-TPC materials; B: loss of methylation in the promoter region results in an increase in gene expression in the high-TPC materials compared to the low-TPC materials.

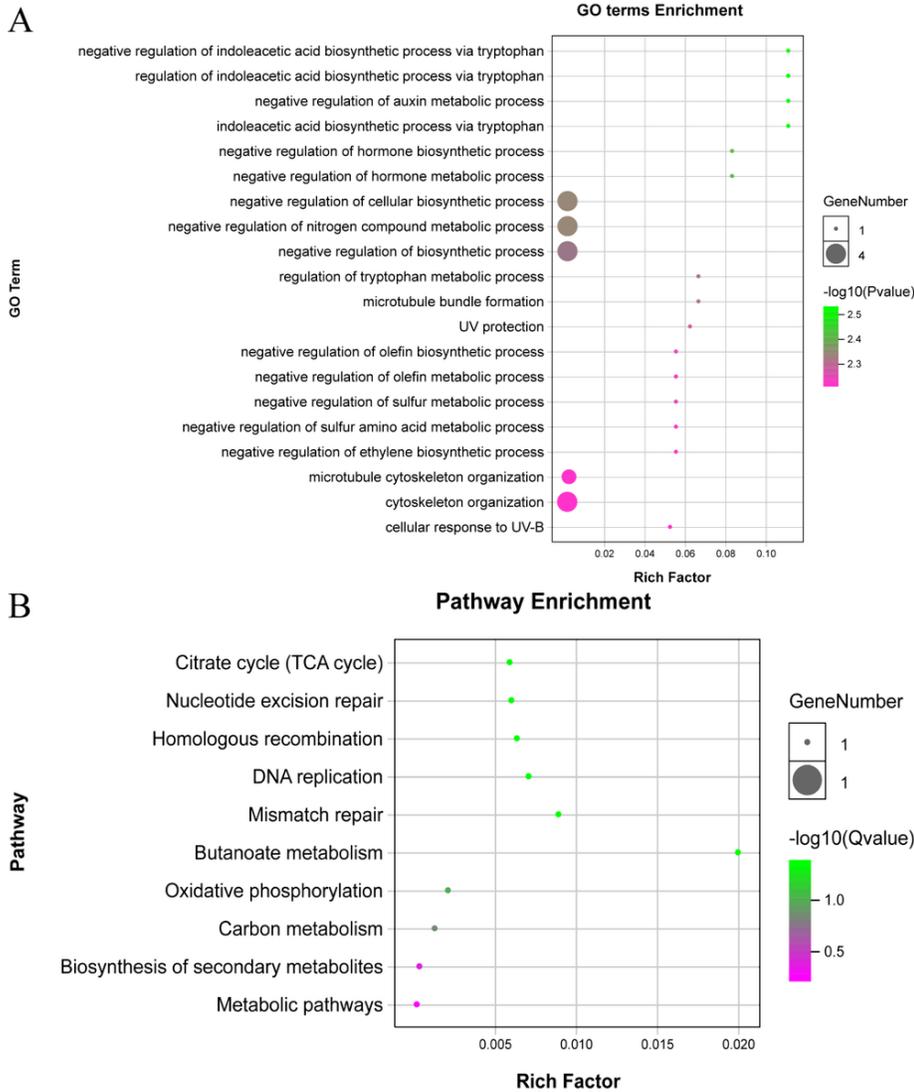


Figure 12

Analysis of GO and KEGG of genes regulated by promoter methylation in two sets of TPC materials

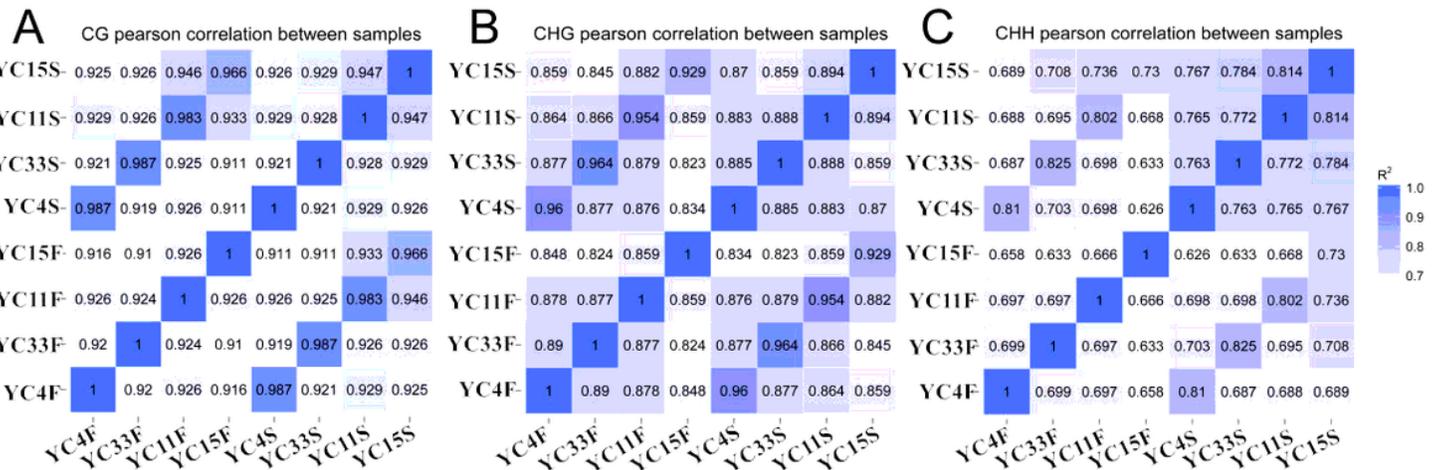


Figure 13

Correlation analysis of three sequences methylation of each sample

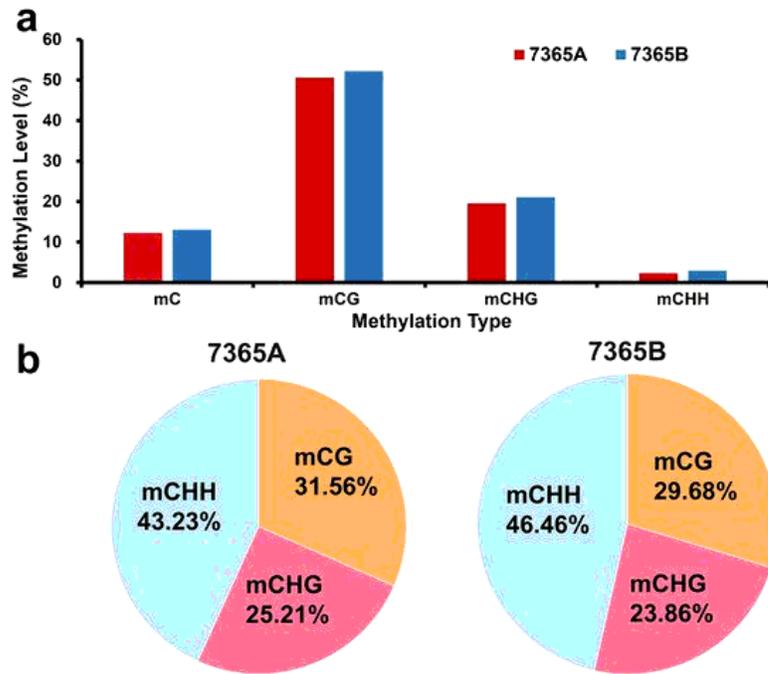


Figure 14

Genome-wide DNA methylation patterns of 7365A and 7365B. (a) Average methylation levels of CG, CHG and CHH in the whole genome; (b) Percentages of methylated cytosines under each context. (Wang et al. 2018b)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterials.docx](#)