



# Transgressive and parental dominant gene expression and cytosine methylation during seed development in *Brassica napus* hybrids

Mauricio Orantes-Bonilla<sup>1</sup> · Hao Wang<sup>2</sup> · Huey Tyng Lee<sup>1</sup> · Agnieszka A. Golicz<sup>1</sup> · Dandan Hu<sup>2</sup> · Wenwen Li<sup>2</sup> · Jun Zou<sup>2</sup> · Rod J. Snowdon<sup>1</sup>

Received: 3 November 2022 / Accepted: 12 March 2023 / Published online: 18 April 2023  
© The Author(s) 2023

## Abstract

**Key message** Transcriptomic and epigenomic profiling of gene expression and small RNAs during seed and seedling development reveals expression and methylation dominance levels with implications on early stage heterosis in oilseed rape.

**Abstract** The enhanced performance of hybrids through heterosis remains a key aspect in plant breeding; however, the underlying mechanisms are still not fully elucidated. To investigate the potential role of transcriptomic and epigenomic patterns in early expression of hybrid vigor, we investigated gene expression, small RNA abundance and genome-wide methylation in hybrids from two distant *Brassica napus* ecotypes during seed and seedling developmental stages using next-generation sequencing. A total of 31117, 344, 36229 and 7399 differentially expressed genes, microRNAs, small interfering RNAs and differentially methylated regions were identified, respectively. Approximately 70% of the differentially expressed or methylated features displayed parental dominance levels where the hybrid followed the same patterns as the parents. Via gene ontology enrichment and microRNA-target association analyses during seed development, we found copies of reproductive, developmental and meiotic genes with transgressive and paternal dominance patterns. Interestingly, maternal dominance was more prominent in hypermethylated and downregulated features during seed formation, contrasting to the general maternal gamete demethylation reported during gametogenesis in angiosperms. Associations between methylation and gene expression allowed identification of putative epialleles with diverse pivotal biological functions during seed formation. Furthermore, most differentially methylated regions, differentially expressed siRNAs and transposable elements were in regions that flanked genes without differential expression. This suggests that differential expression and methylation of epigenomic features may help maintain expression of pivotal genes in a hybrid context. Differential expression and methylation patterns during seed formation in an  $F_1$  hybrid provide novel insights into genes and mechanisms with potential roles in early heterosis.

---

Communicated by Yiqun Weng.

---

Mauricio Orantes-Bonilla and Hao Wang have contributed equally to this work.

---

✉ Rod J. Snowdon  
Rod.Snowdon@agr.uni-giessen.de

<sup>1</sup> Department of Plant Breeding, Land Use and Nutrition, IFZ Research Centre for Biosystems, Justus Liebig University, Giessen, Germany

<sup>2</sup> National Key Laboratory of Crop Genetic Improvement, College of Plant Science & Technology, Huazhong Agricultural University, Wuhan, People's Republic of China

## Introduction

Heterosis refers to the enhanced performance observed in  $F_1$  hybrids derived from two genetically distant, homozygous parents. The improved performance of hybrids compared to their inbred parents was described as early as the late nineteenth century by Charles Darwin during his studies on maize and other plants (Darwin 1876). Since Shull first coined the term “heterosis” in 1914 (Shull 1914; 1948), this phenomenon has been applied in plant breeding to develop hybrids that outperform their inbred parents. However, despite the success of hybrid breeding in many major crops with selfing/outcrossing mating systems, for example maize, sunflower, tomato, sugarbeet and oilseed rape (Steege et al. 2022), the mechanisms driving heterosis have yet to be fully elucidated. Several factors contributing to heterosis

have been proposed: For extensive reviews see for example Wu et al. (2021) and Mackay et al. (2021). Classical quantitative genetics remains core to understanding heterosis as a product of allele interactions through dominance, overdominance, or epistasis (Fujimoto et al. 2018). Nevertheless, through recent research, the cumulative understanding of molecular biology features has raised the question of how non-genetic and non-genomic features are also associated with heterotic patterns and to which extent. Recent studies suggest that gene networks, allele bias, epigenomic and transcriptomic factors play a key role in heterosis (Wu et al. 2021; Yu et al. 2021).

*Brassica napus* (oilseed rape, canola; AACC,  $2n = 38$ ) is not only an important crop where hybrid breeding has been implemented successfully, but is also a model crop of research interest due to its polyploid nature and phylogenetic proximity to *Arabidopsis thaliana*. Oilseed rape is the second most widely grown oilseed crop and has been the third most important food oil source worldwide in the last decade (FAO 2022). The economic importance of hybrid breeding in oilseed rape is evident from the number and relevance of hybrid varieties in key producing countries like Canada, China and Germany. It is estimated that in 2015–2016, at least 80% of oilseed rape grown in China were hybrid varieties (Bonjean et al. 2016), whereas in Canada, the world's largest producer of spring-type canola, herbicide-tolerant hybrid varieties have contributed to significant yield increases during the past decade (Malla and Brewin 2019; FAO 2022). The highest worldwide yields in winter-type oilseed rape in the past ten years were recorded in Germany (FAO 2022), where the percentage of hybrid cultivars registered in the German National List increased from 74% in 2016 to more than 90% in 2022 (Friedt et al. 2018; BSA 2022). These figures highlight the relevance of heterosis in current oilseed rape/canola breeding worldwide. The economic importance increases the need for a more refined understanding of the underlying molecular mechanisms behind heterosis in *B. napus*.

Technological advances in transcriptomic and epigenomic profiling in recent decades have increased awareness of regulatory and epigenetic factors in crop improvement and hybrid breeding (Scossa et al. 2021; Yang et al. 2021). “Omics” technologies not only help to describe and expand genetic diversity in crop species (Louwaars 2018), but can also contribute to elucidating the role of regulatory and non-coding features in plants (Zanini et al. 2022). Transcriptomic and epigenomic features have been widely used to determine molecular and biological functions associated with improved performance in plant hybrids (Yu et al. 2021). For instance, RNA sequencing (RNA-Seq) data developed through microarrays and next-generation sequencing (NGS) has been used to find differentially expressed genes (DEGs) linked to heterosis during diverse growth stages (Wang

et al. 2017a; Zhu et al. 2020). Small RNAs (sRNAs) derived from endogenous genomic loci or exogenous sources are known to regulate various functions and responses in plants. Classification and characterization of microRNAs (miRNAs) and small interfering RNAs (siRNAs) provide valuable information to investigate regulatory factors involved in modulation of gene and trait expression (Griffiths-Jones et al. 2006; Lunardon et al. 2020). For example, sRNAs have been associated with changes in performance in maize, rice and wheat (Zhang et al. 2014a; Li et al. 2014; Seifert et al. 2018a), while epigenomic features, including chromatin interaction, histone modification and DNA methylation, can cause phenotypical changes without alterations in DNA sequences (Fitz-James and Cavalli 2022). Genome-wide methylation differences in various plant species have been associated with phenotypic consequences (Muyle et al. 2022) and linked to heterosis (Kawanabe et al. 2016; Lauss et al. 2018).

Differential gene expression studies in *B. napus* revealed key genes regulating flowering time, disease resistance and abiotic stress (Wu et al. 2016; Wang et al. 2017b; Jian et al. 2019), while small RNA profiling identified microRNA and siRNA sequences associated with pathogen response, abiotic stress and lipid metabolism in oilseed rape (Wang et al. 2017c; Jian et al. 2018; Martinez Palacios et al. 2019; Regmi et al. 2021). Furthermore, DNA methylation patterns were found to contribute to heat response, DNA repair and fertility in *B. napus* (Li et al. 2016; Ran et al. 2016; Wang et al. 2018; Yin et al. 2021).

Nevertheless, few studies have integrated multiple omics strategies to obtain a detailed scenario of expression and methylation patterns in oilseed rape (Shen et al. 2017; Wang et al. 2018). Interestingly, Shen et al. (2017) found specific expression and methylation patterns associated with heterosis in a commercial *B. napus* hybrid cultivar. Enhanced performance due to heterosis has been mostly evaluated at the genomic level and explained through allele interactions (Fujimoto et al. 2018) and introgressions of genomic regions between genetically and genomically distant parents (Hu et al. 2021a; Quezada-Martinez et al. 2021). Nevertheless, the transcriptomic and epigenomic networks involved in heterosis have not been fully elucidated, and the potential to include information on coding and non-coding features in hybrid breeding has been barely explored.

Comprehensive studies in maize and *Arabidopsis* demonstrated that heterosis can be observed in various developmental stages (van Hulst et al. 2018; Zhou et al. 2019). Heterosis during seed development can contribute directly to grain yield, seed biomass, germination and early vigor (Hochholdinger and Hoecker 2007; Jahnke et al. 2010). Since seed formation is characterized by the merging of parental genomes, parent-specific epigenomic effects and genomic imprinting (Thiemann et al. 2009; Castillo-Bravo

et al. 2022), it is an ideal stage for transcriptomic and epigenomic assessments in relation to heterosis. RNA-Seq and methylation-based studies have dissected putative heterotic loci in embryo and seed developmental stages in hybrids of Arabidopsis (Meyer et al. 2012; Kawanabe et al. 2016; Alonso-Peral et al. 2017; Chen et al. 2022). In *A. thaliana* and maize, early heterosis was associated with increases in cell size and number, seed yield and biomass (Jahnke et al. 2010; Wang et al. 2017a; Zhu et al. 2020). Groszmann et al. (2014) found that the maternal genotype was the major determinant of heterosis at early developmental stages in *A. thaliana*. Seed development is also well characterized for enriched epigenomic mechanisms through methylation and transcriptomic regulation, with pollen cells being hypermethylated and ovule cells demethylated in most plants (Batista and Köhler 2020; Montgomery and Berger 2021). Parental dominance effects are attributed with a key role during seed formation through diverging gamete methylation patterns (Weigel and Colot 2012; Lauss et al. 2018). Moreover, the merging of parental genomes during embryogenesis can cause a genomic shock that can further alter the hybrid transcriptome (Bird et al. 2018).

Diverse studies in parent–offspring trios have compared parental dominant and transgressive gene expression patterns via expression level dominance (ELD) analyses in polyploids including *B. napus* (Yoo et al. 2013; Wu et al. 2018; Li et al. 2020) to elucidate the parental genotype effects on gene expression in interspecific hybrids between the diploid species progenitors. In the present study, we analyze transcriptomic and epigenomic differences during seed and seedling development in homozygous parental lines and their  $F_1$  hybrid from a cross between the winter-type *B. napus* accession Express 617 (Lee et al. 2020) and the semi-winter accession *B. napus* G3D001 (Zou et al. 2018).

Preliminary observations by Hu et al. (2021b) showed significant heterosis in hybrids from Express 617 and newly resynthesized oilseed rape lines, including the Express 617 × G3D001 hybrid (encoded in that publication as genotype number T4\_N22) under environmental conditions in China. Numerous studies during the last decades have also shown the heterotic advantages of crossing genetically distant oilseed rape varieties (Qian et al. 2009; Basunanda et al. 2010; Girke et al. 2012; Hu et al. 2021a). Hence, determining differentially expressed and methylated heterotic features between distant germplasm sources can potentially improve our understanding of the molecular mechanisms of heterosis. The overall aim of the present article is to provide an atlas of transcriptomic and epigenomic features associated with heterosis and contribute to dissecting relevant multiomics loci in oilseed rape as a model crop. For this purpose, mRNA, sRNA and whole-genome bisulfite sequencing were carried in the two parental inbreds and their  $F_1$  hybrid. Differential features

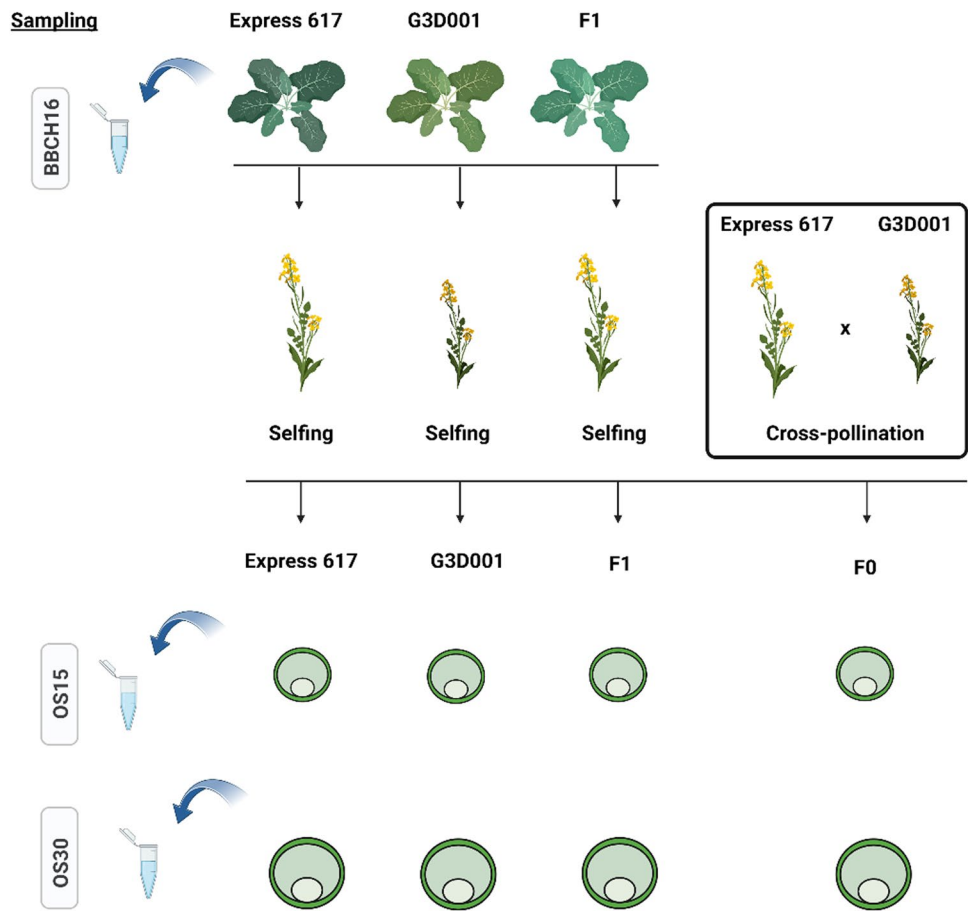
were identified and classified by their respective expression or methylation dominance levels to detect parental and hybrid-specific patterns associated with early developmental stages. Gene ontology enrichment (GO) and integration of omics features were performed to find putative interactions between the detected features and consequently evaluate their epigenomic and transcriptomic impact on early heterosis.

## Material and methods

### Experimental design and growing conditions

Seeds from homozygous, advanced inbred lines of winter-type oilseed rape Express 617 (maternal line), semi-winter semi-synthetic oilseed rape G3D001 (paternal line) and their  $F_1$  hybrid offspring were planted in the 2020–2021 growing season at Huazhong Agricultural University of Wuhan field station. The third youngest leaf from each genotype were sampled from seedlings having six unfolded leaves (BBCH16) at 10:00 am under liquid nitrogen. Flower buds with similar sizes were selected on the fifth day after reaching full flowering (BBCH65) to perform manual selfing on a defined day in all genotypes along with crosses between Express 617 (female pollen recipient) and G3D001 (male pollen donor). The newly generated  $F_1$  crosses were employed to analyze the transcriptomic and epigenomic differences during seed formation between ovules pollinated from selfed  $F_1$  plants and those pollinated by outcrossing from Express 617 and G3D001 (referred to hereinafter as  $F_0$ ). Pollinated ovules were excised with forceps 15 and 30 days after pollination (DAP) at 10:00 am and immediately transferred to liquid nitrogen. Biological replicates consisted of pooled samples from the third youngest leaf from seven individual plants for leaf samples, and from four pollinated ovules from four different plants. The sampled tissue was aliquoted and used for all sequencing types described in this study. Lastly, three independent biological replicates were used for messenger RNA (mRNA) and small RNA expression experiments. Due to low material availability for some samples, only two biological replicates have been used for the methylation studies. Phenotypic measurements were recorded for plant height and dry seed weight for each pooled biological replicate. Significant differences between genotypes were assessed using a one-way analysis of variance (ANOVA) followed by a Tukey test ( $p < 0.05$ ). The experimental designs are summarized in Figs. 1 and S1.

**Fig. 1** Transcriptomic and epigenomic experimental design. Leaves samples were taken at the six-leaves stages (BBCH16) from each biological replicate. Homozygous inbred plants of Express 617 and G3D001 along with their heterozygous  $F_1$  hybrid were self-pollinated to generate selfed ovules from each genotype. The two inbred parents were also crossed during the experiment to develop cross-pollinated ovules ( $F_0$ ). Pollinated ovules were sampled from each biological replicate and sequenced at 15 (OS15) and 30 (OS30) days after pollination



### mRNA, small RNA and whole-genome bisulfite sequencing

mRNA was extracted using TRIzol™ Reagent (Thermo Fisher). A total of 0.5 µg of total RNA per biological replicate were used for preparing 150 bp paired-end (PE) read libraries using the NEBNext® Ultra™ II RNA Library Prep Kit (New England Biolabs, Inc.). Small RNA was extracted using a Plant miRNA kit (Omega Bio-tek Inc.). One microgram of total RNA per biological replicate was employed for the construction of 50 bp single-end (SE) reads using NEBNext® Multiplex Small RNA Library Prep Set for Illumina™ (New England Biolabs, Inc.). Lastly, 2.5 µg of CTAB extracted-DNA per biological replicate were first treated with sodium-bisulfite using the Zymo EZ DNA Methylation-Lightning™ Kit (Zymo Research Corp.) and then built into 150 bp PE read libraries with the TruSeq Nano DNA LT Sample Prep Kit (Illumina Inc.) for whole-genome bisulfite sequencing (WGBS). All libraries were sequenced using an Illumina NovaSeq 6000 platform (Illumina Inc.). Read quality was

evaluated with FastQC v.0.11.9 (Andrews, 2010) and multiqc v.1.9 (Ewels et al. 2016) for all sequencing types. Principal component analyses (PCA) were carried for all libraries per developmental stage using the plotPCA function from DESEQ2 for RNA data (Love et al. 2014) and the *prcomp* R function for WGBS libraries.

### mRNA and sRNA alignments

mRNA libraries were first filtered by selecting reads with an exact length of 150 bp, minimum base quality phred value of 5, no unqualified bases and less than 15% N bases using fastp v.0.23.1 *-q 5 -u 0 -n 15 -l 150* settings (Chen et al. 2018). Splice sites in the Express 617 reference genome assembly (Lee et al. 2020) were identified by first converting the gene annotation file format (Express617\_v1\_gene.gff3; MD5: cf26ec54823f348a0e23f027dc386a16) from a general feature format v.3 (GFF3) to a general transfer format (GTF) using the *agat\_convert\_sp\_gff2gtf.pl* script from AGAT v.0.5.0 (Dainat 2019). The output was then employed to find splice sites with the *hisat2\_extract\_splice\_sites.py* script

from HISAT2 (Kim et al. 2019). An index from the same Express 617 reference was built with *hisat2-build* function, and libraries were then aligned with HISAT2 using the *sensitive* preset and the *known-splicesite-infile* setting with the *hisat2\_extract\_splice\_sites.py* previously generated file as input. Alignments were sorted and converted to a binary alignment map (BAM) format with samtools (Li et al. 2009) *view* and *sort* functions. The number of fragments in genes was counted with featureCounts 2.0.1 (Liao et al. 2014) using the AGAT GTF annotation file and the following settings *-p -B -C -Q 50 -t "exon" -g "gene\_id"*, so that only read pairs having a minimum mapping quality of 50 and both reads aligned to the same strand and chromosome were counted. Genes without any counts in all genotypes were removed. Small RNA libraries were first filtered by removing reads shorter than 18 bp with seqtk v.1.3 (Li 2016). Subsequently, sRNA libraries were aligned against the Express 617 reference using ShortStack v.3.8.5 (Johnson et al. 2016). Only sRNA in which at least 80% of the primary reads had a length between 20–24 nucleotides, with less than 5 unpaired bases in secondary structure, and which were contained in predicted hairpin structures (i.e. only small RNAs clusters with *Y*, *N15*, *N14* or *N13* flags.) were considered as miRNA candidates. Small RNA sequences in which in which 80% of the primary reads had an exact length of 24 nucleotides and without miRNAs selection flags were regarded as putative siRNA based on similar thresholds employed by Lunardon et al. (2020). miRNA and siRNA clusters without any coverage in all biological samples were discarded prior to differential expression analysis.

### Expression level dominance analysis

Differential mRNA, miRNA and siRNA expression patterns between the  $F_1$  hybrid and its parents were assessed by comparing tissues within genotype trios in five tissues/stages: leaves from parental and  $F_1$  plants at six-leaf stage (BBCH16); ovules 15 days after pollination from selfed parents and  $F_1$  (OS15- $F_1$ ) or  $F_0$  (OS15- $F_0$ ); and ovules 30 days after pollination from selfed parents and  $F_1$  (OS30- $F_1$ ) or  $F_0$  (OS30- $F_0$ ), respectively. Differentially expressed genes, differentially expressed miRNAs (DE-miRNAs) and differentially expressed siRNAs (DE-siRNAs) between genotypes for each stage were identified using the counts from each biological replicate with DESEQ2 (Love et al. 2014) with a *padj* value threshold  $< 0.05$ . The DESEQ2 built-in *estimateSizeFactors* and *counts* functions were used to extract the normalized counts which were then used for expression level dominance analyses. Briefly, student's *t* test ( $p < 0.05$ ) from normalized counts of DEGs and DE-miRNAs identified in DESEQ2 were run between all genotypes for each comparison stage and gene. Tukey tests ( $p < 0.05$ ) were then carried to rank

each genotype by expression level. Finally, the resulting patterns were divided based on Yoo et al. (2013) as additive (I, XII), paternal dominant (II, XI), maternal dominant (IV, IX) and transgressively up- (III, VII, X) or downregulated (V, VI, VIII). Gene expression heatmaps were generated with idep93 (Ge et al. 2018) using correlation distances and average linkages, and differentially expressed genes or sRNAs shared between all stages were detected using the *Venn Diagrams* tool (VIB-UGent 2021). In addition, the percentages of upregulated and downregulated DEGs from all genes per subgenome, genotype and stage were calculated to evaluate subgenomic expression bias and normalized gene transcript values were summarized as heatmap values for easier comparison.

### Gene ontology enrichment

Gene models in the Express 617 reference assembly were functionally annotated through synteny comparison against the Darmor v.4.1 genome (Chalhoub et al. 2014) with inparanoid v.4.2 (O'Brien et al. 2005) using bootstrap, a BLOSUM80 (BLOcks SUBstitution Matrix) and an initial cutoff score of 60. Inparalogs with a similarity score equal or greater than 70 were selected for each gene. Pairs with only one homolog and with the highest similarity score were kept. The homologs were used for GO enrichment of biological processes based on expression level dominance for each stage, as well in comparisons between the  $F_1$  and  $F_0$  genotypes, using ShinyGo v.0.76 (Ge et al. 2020) with a 0.05 false discovery rate (FDR) cutoff. Only biological functions with more than one gene per biological pathway and with at least two GO groups were selected.

### DE-miRNA target prediction and mRNA interaction

Differentially expressed miRNAs sequences were extracted and used to predict their corresponding targets in Express 617 gene models using psRNATarget (Dai et al. 2018) with the version 2 scoring schema (Axtell 2013). Maximum unpaired energy (UPE) of 25 and a flank length between 13 to 17 nucleotides in up/downstream region were set as target accessibility cutoffs. All possible targets for DE-miRNAs were reported since each miRNA can have multiple mRNA targets due to isomiRs formation. The DE-miRNAs were classified into putative miRNA families by blasting their sequences with BLAST (Altschul et al. 1990) against the mature miRNAs from the Brassicaceae family available at the miRBase sequence database release version 22.1 (Griffiths-Jones et al. 2006). Only the top five matches with the highest alignment scores and lowest expect values for each DE-miRNA were retained. Stem-loop sequences from the Brassicaceae family were used as BLAST targets when no mature miRNAs matches were found. Alternatively, if

no Brassicaceae matches were found, then mature miRNAs and stem-loop sequences from the Viridiplantae clade were employed. The expression patterns from miRNA targets that were DEGs were compared with their associated targeting DE-miRNA expression to evaluate possible interactions between miRNA and mRNA target. The DEGs target functions were estimated by comparing their coding sequences against the Araport v.11 *Arabidopsis thaliana* coding sequences model (Cheng et al. 2017) via BLAST. Only the hit with the lowest expect value and not greater than  $1.0 \times 10^{-4}$ , lowest identity percentage equal or above 90% and without gaps were selected.

### Bisulfite sequencing alignment and methylation level dominance

Reads with a minimum base quality phred value of 5, unqualified base percent limit of 50 and less than 15% N bases were selected from WGBS libraries using fastp v.0.23.1 *-q 5 -u 50 -n 15* settings (Chen et al. 2018). TrimGalore (Krueger et al. 2021) was then employed for trimming 8 basepairs from both 5' and 3' ends for each library as recommended for TruSeq libraries in the Bismark documentation. The Express 617 reference genome was bisulfite converted and indexed with Bismark v.0.23 (Krueger and Andrews 2011) *bismark\_genome\_preparation* tool. Filtered reads for each biological replicate were aligned to the bisulfite converted genome using *bismark* under default settings. Duplicates were afterward removed with *deduplicate\_bismark* and methylated cytosines (mC) were extracted using *bismark\_methylation\_extractor* while ignoring the first 2 basepairs from both 5' and 3' ends for both reads of a pair. Every mC in a CpG, CHG or CHH methylation context was selected and converted to a browser extensible data (BED) format with *bismark2bedGraph* using the *-cutoff 3 -CX -and -scaffolds* settings to select all nucleotides in which the methylation state was reported at least thrice.

The coverage for each mC in every methylation context per biological replicate was calculated with the *coverage2cytosine* from the Bismark package. The mC coverage in assigned chromosomes was then used as input for DMRCaller v. 1.22.0 (Catoni et al. 2018) to detect differentially methylated regions (DMRs). Each genotype within a trio was compared to each other using the *computeDMRs* function in 1000 bp bins with the *bins* method and the following settings: score test, a 0.01 *p* value threshold, and minimum cytosine count, methylation proportion difference and gap between bins of 4, 0.4 and 0 accordingly. The DMR methylation levels (i.e. the number of reads supporting methylation) were extracted from DMR output files and student's *t* test ( $p < 0.05$ ) was run between all genotypes for each stage and DMR. Tukey tests ( $p < 0.05$ )

were then used to rank the methylation within DMRs between genotypes and classified them by methylation level dominance (MLD) following the same categorization employed for ELD by Yoo et al. (2013) as additive (I, XII), paternal dominant (II, XI), maternal dominant (IV, IX) and transgressively hyper- (III, VII, X) or hypomethylated (V, VI, VIII). Shared and unique DMR across all stages were found with the *Venn Diagrams* tools (VIB-UGent 2021). Moreover, heatmaps displaying the genome-wide methylation levels between biological replicates per methylation context in 100 kbp bins were generated using the *circlize* and *ComplexHeatmap* packages (Gu et al. 2014, 2016).

### Cytosine methylation statistics and identification of methylated features

The number of mC nucleotides and the cytosine methylation level per 1 kbp bin (i.e., numbers of reads supporting cytosine methylation per bin) in each methylation context, genotype and stage were determined based on Bismark's *coverage2cytosine* generated files using *bedtools makewindows* and *intersect* functions (Quinlan and Hall 2010). In addition, DMRs were intersected with exons, introns, repeats and 1 kbp upstream promoter regions from Express 617 using *bedtools intersect* function. GO enrichment was analyzed for differentially expressed genes having DMRs for all stages and genotypes. If no enrichment was detected, then the most frequent biological functions found in Ensembl Biomart (Cunningham et al. 2022) *B. napus* reference (Chalhoub et al. 2014) were reported. Detected differentially expressed genes having an additive or dominant expression level dominance pattern were defined as putative genetic epialleles if their loci coincided with corresponding additive or dominant methylation patterns in DMRs. Furthermore, the correlation between gene expression and methylation in either the gene body or promoter were assessed with a Kendall  $\tau$  test using the *cor.test* function in *R*, since the pre-evaluation of our gene expression data showed a non-Gaussian distribution, as also observed in other transcriptomic studies (Robinson et al. 2010; Di et al. 2011; Love et al. 2014; Church et al. 2019).

Heatmaps comparing the gene methylation in gene bodies and promoters and gene expression in transgressive DEGs were made with Heatmapper (Babicki et al. 2016) using Euclidean distances and average linkages to analyze the putative interaction between expression and methylation. Moreover, repeats in the Express 617 assembly were assigned to repeat families using RepeatModeler (Smit and Hubley 2008) and CpG islands were identified with *cpGplot* from the EMBOSS v.6.6.0 package (Rice et al. 2000). CpG islands were called if the GC% was equal or greater

than 50%, length greater than 200 bp and a minimum 0.6 observed to expected CpG dinucleotides ratio as described by Gardiner-Garden and Frommer (1987). Additionally, plots showing DEGs and methylation levels for each chromosome and stage, centromere loci and repeat density were made using the *circlize* package (Gu et al. 2014). Repeat density for each 1 kbp bin within each chromosome was calculated using *bedtools* while predicted Express 617 centromere loci were added based on Orantes-Bonilla et al. (2022).

Lastly, DMRs were intersected with DE-siRNAs, CpG islands and transposable elements (TEs) in 5 kbp upstream and downstream gene and DEG-flanking regions in assigned chromosomes using *bedtools* to evaluate putative interactions between differentially methylated features and gene expression during seed development. The threshold was selected based on previous work on transposable elements and genomic imprinting in *B.napus* by Rong et al. (2021) and the fact that the average distance between genes in assigned chromosomes of the Express 617 reference is approximately 7.5 kbp. Chi-square tests followed by an FDR post-hoc adjustment ( $p < 0.05$ ) were performed to find significant associations between differentially methylated and non-methylated features and determine their respective distances to genes or DEGs across all stages.

### Segmental expression assessment

Clustering of DEGs across chromosomal segments observed on *circlize* generated plots was further investigated. To assess the presence of expression clusters, segments that had more than 20 DEGs over a 500 kbp window were considered as putative differentially expressed segments. The threshold was selected on the basis that the Express 617 genome assembly has an average of 200 genes per 500 kbp and hence 20 genes would correspond to 10% of genes in the segment. The ratio of upregulated to downregulated DEGs per genotype and stage in each segment was calculated and normalized to Z-scores. Only segments showing clear differential patterns between genotypes based on Z-score heatmap clustering were retained. Such segments could either be a result of parental expression bias or due to commonly observed genomic rearrangements in allopolyploid *B. napus*. To investigate both possibilities, short read genomic sequence data from a G3D001 biological replicate was used for calling copy number variation (CNV) and investigating putative linkages between structural rearrangements and expression patterns. For this purpose, genomic DNA from a G3D001 ovule biological replicate taken 30 days after pollination was extracted using a CTAB protocol (Doyle and Doyle 1987). Paired-end libraries were built with KAPA HyperPlus Kit (KAPA Biosystems) and sequenced with an Illumina NovaSeq 6000 platform (Illumina Inc.). Read quality was

evaluated with FastQC v.0.11.9 and libraries were afterward aligned with *minimap2* (Li 2018) against the Express 617 genomic reference (Lee et al. 2020). Alignments with both forward and reverse reads properly mapped (flags 99, 163, 147 and 83) were selected with *samtools view* and used to calculate coverage across chromosomes using the *bamtobed* and *genomecov* functions from *bedtools*. The coverage was used as input in a previously described deletion-duplication pipeline (Stein et al. 2017), modified by excluding outliers if the depth was above 100 and by defining deletions and duplications as 25 kbp length segments that are one standard deviation above or below the mean coverage. Deletions and duplications were recorded in tab-separated files and intersected with differentially expressed segments using the *intersect* function in *bedtools*.

## Results

### Maternal dominant expression and methylation increases during seed development

Gene and small RNA expression as well as genome-wide methylation patterns from Express 617, G3D001 and their hybrid were compared during seed and seedling developmental stages. The parents were crossed during the experiment to evaluate developmental differences between selfed- $F_1$  plants and Express 617  $\times$  G3D001 pollinated ovules that would develop into  $F_1$  plants (referred here as  $F_0$ ). Figure 1 provides an overview of the plant materials, sampling tissues/timepoints and sample nomenclature and Fig. S1 summarizes the bioinformatic pipeline used. Next-generation sequencing yielded abundant coverage for each biological replicate (Tables S1, S2, S3). Approximately 6.8 Gbp of mRNA sequences per biological replicate were aligned against the Express 617 genome assembly using HISAT2 v. 2.2.1 splice site aware aligner (Kim et al. 2019), producing mean alignment rates of 98.2% (Table S1). In addition, an average of 31 million sRNA reads per biological replicate were used to find putative miRNA and siRNA sequences with ShortStack v.3.8.5 (Johnson et al. 2016). Overall, each sRNA cluster had an average coverage depth of 186 (Table S2). Moreover, whole-genome bisulfite treated reads having a 31  $\times$  genome coverage per biological replicate were aligned and processed with Bismark v.0.23 (Krueger and Andrews 2011), as reported in Table S3. PCA plots from mRNA and sRNA libraries showed an overall clustering of biological replicates (Fig. S2, S3, S4). PCA plots from methylation levels displayed grouping of biological replicates to a slightly lesser extent since methylation level is calculated in the whole genome unlike RNA and sRNA counts that are only derived from a specific set of features. Nonetheless, genome-wide methylation level heatmaps

suggested an overall agreement of methylation level between biological replicates for each stage and context (Fig. S5, S6, S7, S8, S9, S10).

All alignments were then employed to find features that were differentially expressed or differentially methylated between genotypes across all stages. In summary, a total of 31,117 DEGs, 344 DE-miRNAs, 36,229 DE-siRNAs and 7399 DMRs in both CpG and CHG methylation contexts were identified across all possible parents-hybrid comparisons per stage (Tables S4, S5, S6, S7, S8, S9). The detected features were evenly distributed across all chromosomes (Tables S10, S11, S12, S13, S14). Differential features were further classified by their expression and methylation level dominance (Fig. 2). More than 90% of the differentially expressed and methylated features corresponded to parental dominant (II, XI, IV, IX) and additivity (I, XII) models based on the dominance classification proposed by Yoo et al. (2013). Moreover, maternal dominance accounted for approximately 89%, 85%, 83% and 60% from all detected DEGs, DE-miRNAs, DE-siRNAs and DMRs in the  $F_0$ ,

respectively, whereas paternal dominance was more prevalent in the  $F_1$ -selfed offspring (Table S15). Furthermore, most maternal dominant DMRs in the  $F_0$  were hypermethylated, whereas DEGs were downregulated. This contrasts to the expected female gamete demethylation observed in seed formation in other plants (Batista and Köhler 2020).

Transgressive upregulated features, in which the hybrid has a higher expression than the parents, were more frequent in seeds from selfed- $F_1$  plants compared to those from the recently formed  $F_0$ . Maternal dominance from Express 617 accounts for most of the DEG and DE-siRNAs patterns observed, suggesting a potential maternal relevance in seed development. Interestingly, no gene expression bias was found between the A and C subgenomes (Table S16, Figure S11, S12, S13, S14, S15); nevertheless, more upregulation was observed in the paternal line, while the maternal line displayed more downregulation during seed development. This contrasts with the expected gene silencing in the maternal genome that is attributed to maternal demethylation during seed formation. Moreover, a slightly higher number of

Feature	Stage	Additivity		G3D001 dominance		Express 617 dominance		Transgressive-down regulation/hypomethylation			Transgressive up regulation/hypermethylation		
		I	XII	II	XI	IV	IX	III	VII	X	V	VI	VIII
		E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G
DEG	BBCH16	25.19	8.68	9.01	20.34	15.05	18.75	0.08	0.77	0.19	0.13	0.18	1.62
	OS15-F1	9.79	6.44	9.59	18.16	13.20	24.58	0.89	3.64	0.26	0.49	1.53	11.43
	OS15-F0	0.55	0.08	0.09	1.04	46.34	50.77	0.09	0.13	0.00	0.00	0.50	0.40
	OS30-F1	8.75	5.44	5.49	11.71	20.68	37.93	0.81	1.73	0.06	0.28	2.32	4.81
	OS30-F0	7.37	2.18	1.93	3.21	32.68	48.52	0.34	0.53	0.01	0.05	1.38	1.81
DE-miRNA	BBCH16	13.85	9.23	9.23	27.69	9.23	20.00	0.00	9.23	0.00	1.54	0.00	0.00
	OS15-F1	5.38	5.38	8.60	23.66	18.28	23.66	0.00	0.00	0.00	3.23	0.00	11.83
	OS15-F0	1.96	0.00	0.00	2.94	42.16	50.00	0.00	0.00	0.00	0.00	1.96	0.98
	OS30-F1	14.74	1.99	2.39	8.37	23.90	45.02	0.40	0.40	0.00	0.00	0.40	2.39
	OS30-F0	14.12	1.15	0.76	3.44	33.21	45.04	0.38	0.38	0.00	0.00	0.00	1.53
DE-siRNA	BBCH16	16.30	7.54	8.83	30.81	15.39	19.82	0.05	0.56	0.12	0.02	0.12	0.44
	OS15-F1	6.69	8.67	16.20	25.78	11.78	17.36	0.03	0.26	0.01	2.50	2.46	8.25
	OS15-F0	1.06	0.16	0.09	1.30	44.77	49.02	0.10	0.11	0.00	0.00	2.16	1.23
	OS30-F1	12.58	2.58	3.58	13.20	23.07	40.56	0.27	0.76	0.05	0.16	0.60	2.60
	OS30-F0	12.91	1.46	2.12	7.73	28.63	44.42	0.19	0.50	0.04	0.08	0.48	1.46
DMR	BBCH16	42.94	26.46	5.51	13.31	11.11	0.00	0.00	0.00	0.03	0.06	0.39	0.19
	OS15-F1	49.10	9.16	13.21	9.33	17.74	0.00	0.00	0.00	0.03	0.49	0.55	0.39
	OS15-F0	5.08	2.29	0.00	0.15	46.47	15.04	0.03	0.00	0.00	0.03	29.37	1.55
	OS30-F1	39.17	14.23	15.77	8.24	21.13	0.00	0.00	0.00	0.49	0.52	0.44	0.44
	OS30-F0	21.99	14.88	2.21	0.69	49.12	8.99	0.00	0.00	0.00	0.03	2.03	0.06

**Fig. 2** Percentages of differentially expressed genes (DEGs), differentially expressed miRNAs (DE-miRNAs), differentially expressed siRNAs (DE-siRNAs) and differentially methylated regions (DMRs) in CpG and CHG methylation contexts by expression level dominance (ELD) and methylation level dominance (MLD) patterns per stage. Increase and decrease in expression and methylation per pattern are displayed by the dot-ended lines showing the relative expression or methylation levels for the parental genotypes Express 617 (E) and

G3D001 (G) along with their  $F_1$  hybrid (F). Differential expression and methylation are displayed for leaf samples at stage BBCH16 and for ovules at 15 (OS15) and 30 (OS30) days after pollination by selfing ( $F_1$  ovules) or cross-pollination between the two parental lines ( $F_0$  ovules). Percentages are displayed with colored backgrounds to represent high (red) or low (blue) abundance

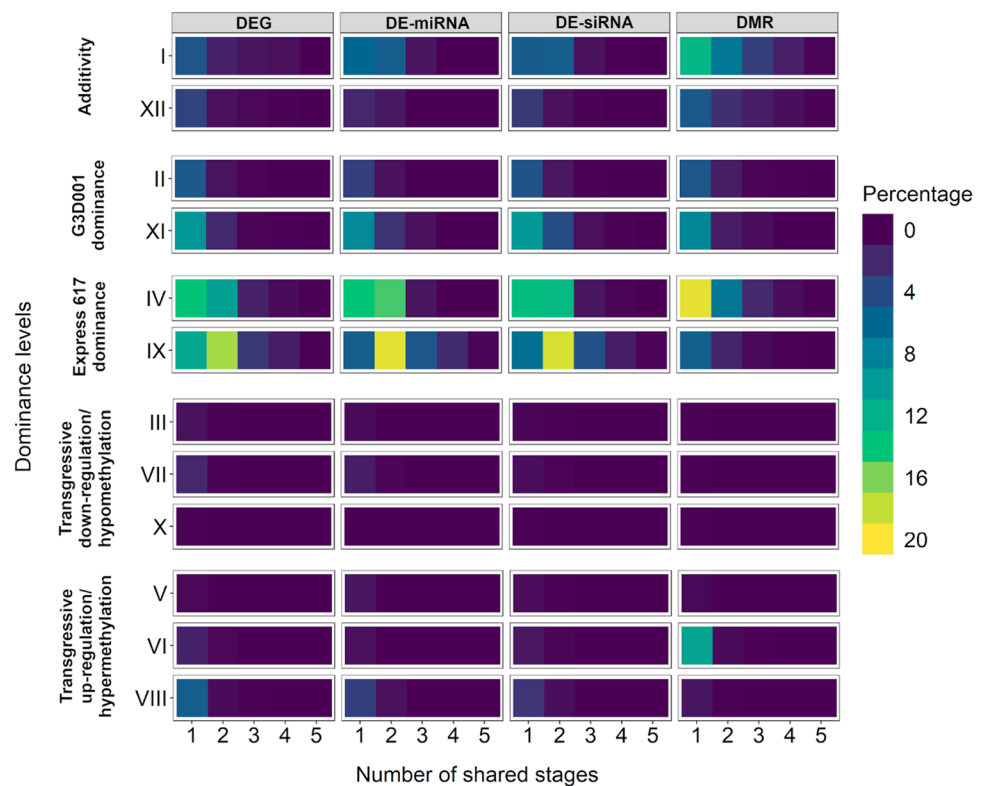
differentially expressed features following maternal expression patterns were found in the  $F_0$  than in the selfed- $F_1$ ; however, this might be due to the allele segregation in the selfed- $F_1$  plants that would lead to the maternal parent being heterozygote and putatively reducing the number of features with maternal dominant expression.

A total of 1565 DEGs, 12 DE-miRNAs, 1111 DE-siRNAs, 896 DMRs in CpG context and 650 DMRs in CHG context were consistently detected across all stages (Tables S17). Altogether, differential features present consistently in all stages and genotypes corresponded to 3% of all detected features, whereas differential features unique to a single stage compromised approximately 2% of all detected features (Table S17). Differential features with consistent dominance level patterns across sampling stages are presented in Fig. 3. Interestingly, features with consistent dominance patterns among stages tended to exhibit maternal dominance. These features were mostly shared between early and late pollinated ovule stages in the  $F_1$  and  $F_0$  (Table S18, S19, S20, S21), suggesting dominance of the maternal genotype during seed formation for both genotypes. Parental effects have been reported to play a role in heterosis in maize and Arabidopsis (Ma et al. 2018; Castillo-Bravo et al. 2022) and are further addressed in the discussion.

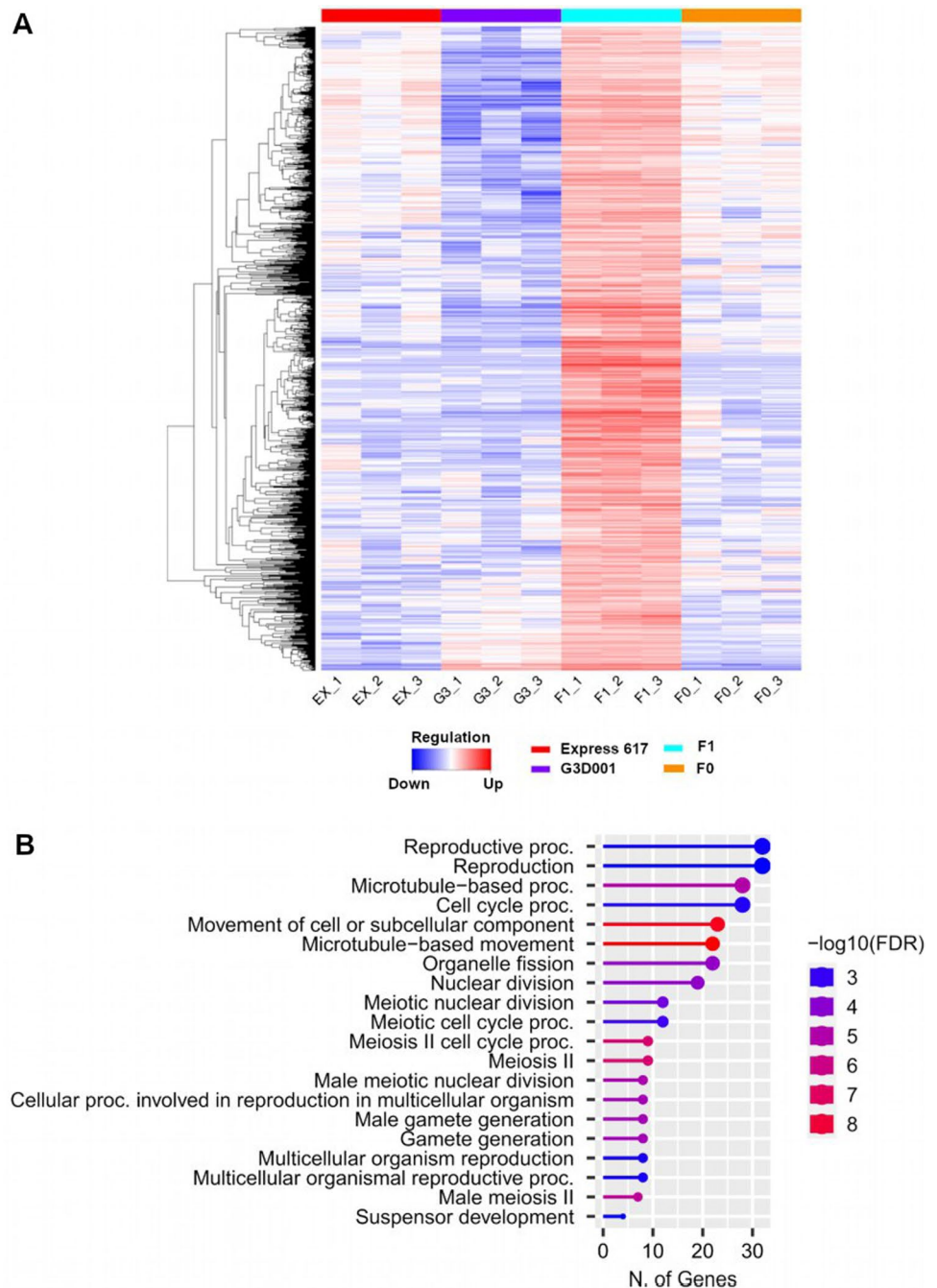
## DEGs and differential miRNA expression regulate $F_1$ seed development

Gene ontology enrichment for biological processes was carried out for all stages based on their expression level dominance. Significant enrichment for pivotal biological functions such as amino acid and carbohydrate synthesis, stress response, photosynthesis, protein transport and DNA repair and replication were found in 15 and 30 days after pollination ovules (Table S22). No significant enrichment was identified in leaves during the seedling stage. Only transgressively upregulated genes in  $F_1$  ovules after 15 days of pollination displayed terms associated with reproduction and meiosis (Fig. 4, Table S23, Figures S16-S20). Differential gene expression and gene ontology between the  $F_1$  and  $F_0$  at 15 days after pollination showed an increase in photosynthesis-related functions in the  $F_1$  hybrid, whereas the  $F_0$  showed increased accumulation of energy reserve compounds and cell mobilization (Table S22). GO terms linked to carbohydrate metabolism, photosynthesis, stress response and cell development have been linked to heterosis in maize, rice, sunflower and oilseed rape (Bao et al. 2005; Lai et al. 2006; Ma et al. 2018; Zhu et al. 2020). Phenotypic measurements recorded during the growing of the parents and the hybrid genotypes used in this study revealed that hybrids had a more robust plant architecture and higher dry seed weight than their parents (Table S24, Fig. 5), in concordance, with

**Fig. 3** Percentage of shared differential features between stages based on dominance level patterns displaying differentially expressed genes (DEGs), differentially expressed miRNAs (DE-miRNAs), differentially expressed siRNAs (DE-siRNAs) and differentially methylated regions (DMRs) in CpG and CHG contexts



**Fig. 4** (a) Gene expression heatmap and (b) gene ontology (GO) enrichment of biological processes from 15 days after pollination ovules with transgressive upregulation patterns in the  $F_1$

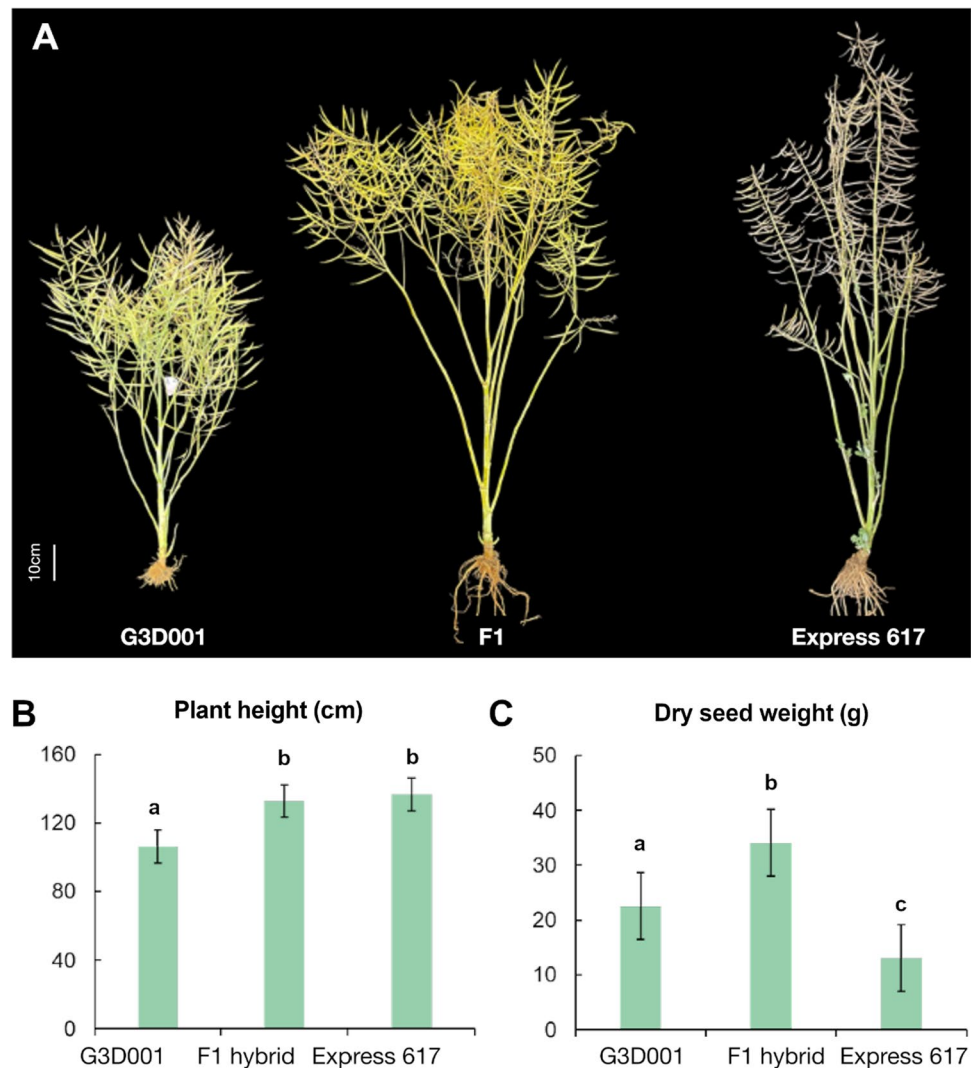


previous observations on this and other related hybrids (Hu et al. 2021b; Orantes-Bonilla et al. 2022).

Additionally, 51 putative mRNA targets from all DE-miRNAs were detected across all stages (Table S25). Interactions between DE-miRNAs and DEG mRNA targets are reported in Table 1 and Table S26. Most DE-miRNAs associated with DEG targets had downregulated expression in the parents and  $F_1$  (ELD group IX) and were more abundant during the late seed developmental stage. Expression from *B. napus* orthologs of *PHABULOSA* (*PHB*), *REVOLUTA* (*REV*) and *TARGET OF EARLY ACTIVATION TAGGED*

2 (*TOE2*), which are involved in plant growth and development, was not increased despite the low expression of miRNAs known to target these genes. Likewise, positive proportional expression interactions were observed in a *B. napus* ortholog of *EMBRYO DEFECTIVE 2204* (*EMB2204*) on chromosome A02, whereas an inversely proportional interaction between the miRNA and mRNA target was found for a *B. napus* ortholog of *EMBRYO DEFECTIVE 2016* (*EMB2016*) on chromosome A03 (Fig. 6). Both genes are involved in embryo development, yet they appear to be regulated in an opposite manner, although further research

**Fig. 5** Phenotypes for field-grown plants of inbred *B. napus* parents G3D001, Express 617 and their  $F_1$  hybrid showing (a) plant architecture, (b) plant height and (c) dry seed weight during an experimental trial in Wuhan, Central China. Biological replicate averages and standard deviations are shown in (b) and (c) for plant height and dry seed weight, and genotypes showing significant differences ( $p < 0.05$ ) detected by Tukey tests are indicated with letters above the bars



is required to elucidate their role in *B. napus* seed formation. *PHB*, and possibly *PHAVOLUTA* (*PHV*), are positive regulators of the *LEAFY COTYLEDON 2* (*LEC2*), a well-known regulator of seed maturation (Tang et al. 2012). A BLAST search from the *A. thaliana* Araport 11 assembly (Cheng et al. 2017) for the *LEC2* coding sequence (*AT1G28300.1*) revealed a single hit that passed the filtering criteria specified for miRNA targets in Materials and Methods. The ortholog corresponded to the *C05p022870.1\_BnaEXP* gene model in the Express 617 genome assembly (Lee et al. 2020), which was found to be differentially expressed in late seed development (Table S18 and S26). This example further highlights the broader and indirect impact of miRNAs through gene network interactions. Such transcriptomic networks not only play critical roles in heterosis (Wu et al. 2021) but are also regulated partially by miRNAs (Dong et al. 2022).

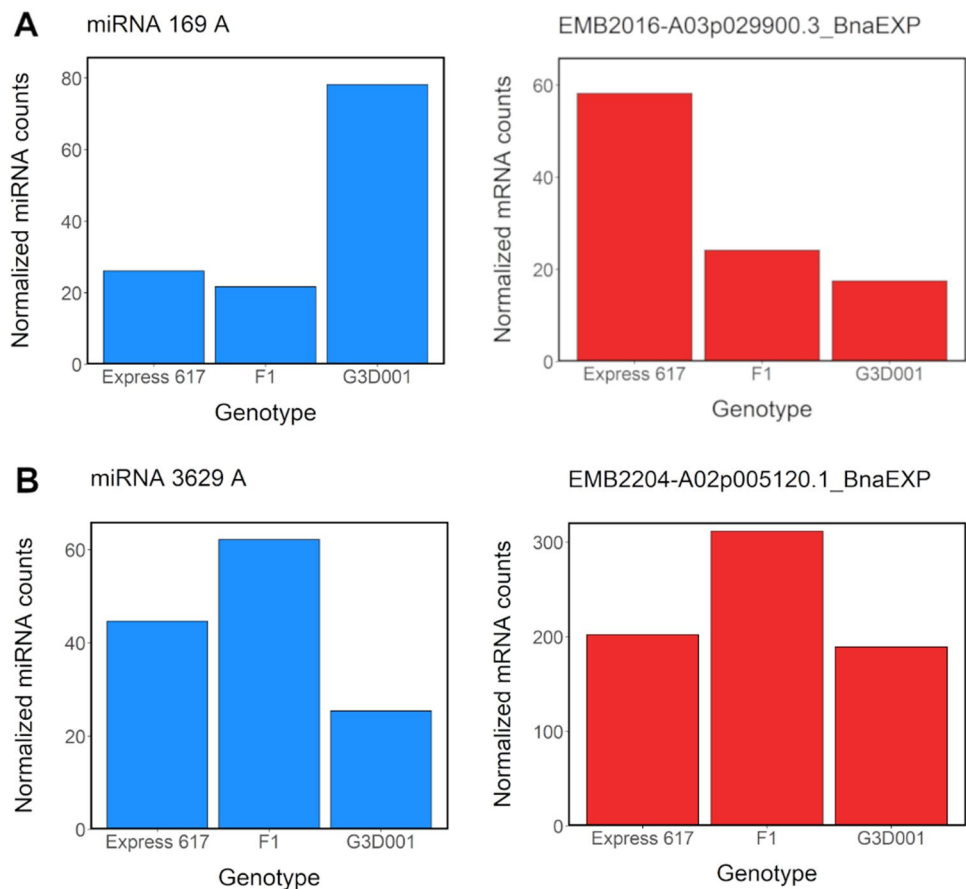
### Methylated features in early seed formation

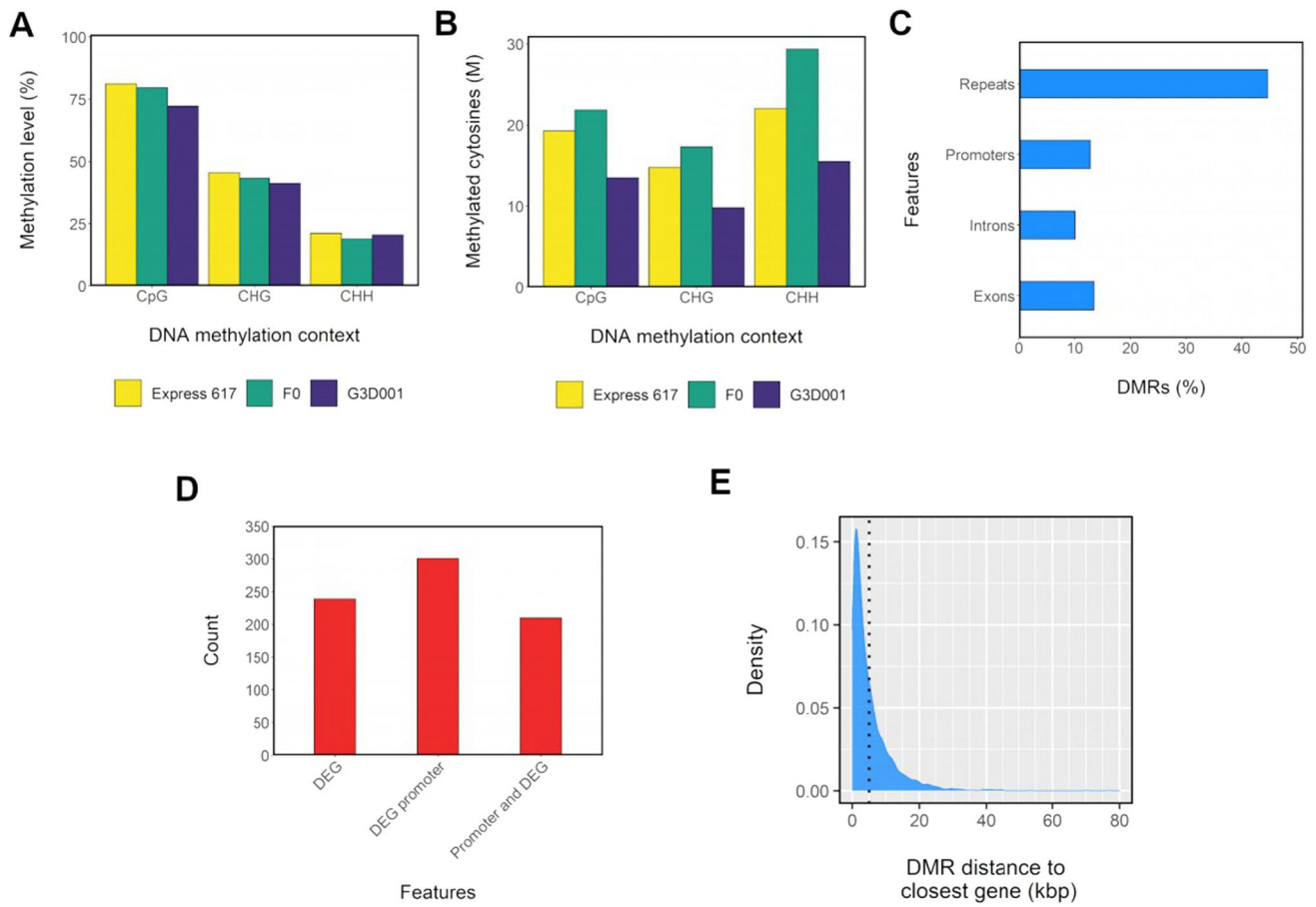
Methylation levels were highest in the CpG context, with an average of 80% across all stages and genotypes (Fig. 7). Methylation levels in CHH context were the lowest, ranging from 20 to 27% despite having the highest number of methylated cytosines (Table S27, Fig S21, S22, S23, S24). No methylation bias per chromosome was observed (Table S27). Approximately 12%, 14% and 10% of DMRs were in promoters, exons and introns, respectively, whereas a high percentage of DMRs (43%) were inside repeat motifs (Table S28). Repetitive sequences account for 37.5% of the Express 617 genome (Lee et al. 2020), and although 66% of repeats were methylated with an average 41% methylation level, less than 1% were differentially methylated (Table S29). Most differentially methylated transposable element (TE) families and superfamilies coincided with those that are most frequent in the Express 617 reference genome, such as LTR (long terminal repeat) Copia and Gypsy families. Approximately 70%

**Table 1** Predicted mRNA target from differentially expressed miRNAs (DE-miRNAs) in 15 and 30 days after pollination ovules in  $F_1$  and parents by expression level dominance (ELD)

Stage	Predicted DE-miRNA family	DE-miRNA-ELD	miRNA Target	miRNA target-ELD	<i>A. thaliana</i> homolog ID	<i>A. thaliana</i> homolog name
OS15- $F_1$	miRNA 165/166 A miRNA 165/166 B miRNA 165/166 C	IX	C04p041520.1_BnaEXP	IX	AT2G34710	PHABULOSA
OS30- $F_1$	miRNA 3629 A miRNA 166 A miRNA 165/166 B miRNA 166 B miRNA 166 C miRNA 166 A miRNA 165/166 B miRNA 166 B miRNA 166 C miRNA 9410/9411 A miRNA 9410 A miRNA 165/166 B miRNA 169 A miRNA 172 A	VIII IX  IX  IX  IX  IX  IX  IX  IX  IX	A02p005120.1_BnaEXP C04p041520.1_BnaEXP  C05p024820.1_BnaEXP  A06p017700.1_BnaEXP  A02p009610.1_BnaEXP A03p029900.3_BnaEXP C09p033660.1_BnaEXP	VIII IX  XII  VII  IX XI IX	AT1G22090 AT2G34710  AT1G30490  AT3G47060  AT5G60690 AT3G05680 AT5G60120	EMB2204 PHABULOSA  PHAVOLUTA  FTSH PROTEASE 7  REVOLUTA EMB2016 TOE2

**Fig. 6** Normalized expression levels from selected differentially expressed miRNAs (DE-miRNA) and their respective differentially expressed target genes (DEG) in ovules 30 days after pollination in the  $F_1$  and parental genotypes, respectively. **(a)** Inversely proportional miRNA-mRNA target expression of miRNA 169A and a *B. napus* ortholog of its target gene *EMB2016* on chromosome A03 (A03p029900.3\_BnaEXP). **(b)** Proportional miRNA-mRNA target expression of miRNA 3629A and a *B. napus* ortholog of its target gene *EMB2204* on chromosome A03 (A02p005120.1\_BnaEXP)





**Fig. 7** Methylation patterns in 15 days after pollination ovules from  $F_0$  and parents. **(a)** Methylation level per genotype and DNA methylation context. **(b)** Count of methylated cytosines in million ( $M$ ) scale per genotype and DNA methylation context. **(c)** Distribution of differentially methylated regions (DMRs) across introns, exons, repeats

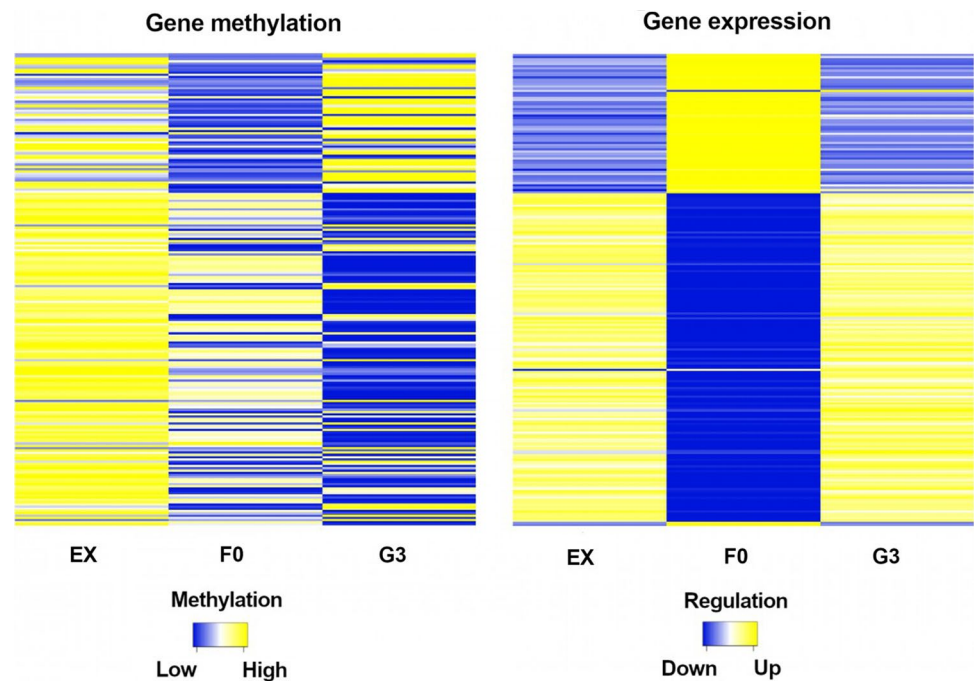
and promoters (1 kbp upstream from gene start). **(d)** Distribution of methylated differential expressed genes (DEGs) and their promoters. **(e)** Kernel density estimation (KED)-based distribution of DMRs distance to closest gene. A dotted line is used to delimit DMRs located 5 kbp from a gene

of these were found within 5 kbp flanking regions of genes (Table S30, S31). Chi-square tests followed by FDR adjusted post-hoc testing at  $p > 0.05$  revealed a significant association between the analyzed genomic features (DMRs, TEs, and differentially methylated or non-methylated DE-siRNAs) in terms of their distance to genes and DEGs (Table S32). Interestingly, around 70% of the detected features were in 5 kbp gene flanking regions; nevertheless, only less than 20% of them were found in 5 kbp DEG-flanking regions (Table S32). Moreover, most DMRs inside gene flanking regions converged more around non-DEGs in comparison to DEGs (Tables S32–S33), which suggests an overall conserved pattern of gene regulation despite differential methylation across genotypes.

A total of 392 genes that were both differentially expressed and differentially methylated were regarded as putative epialleles (Table S34). No gene ontology enrichment was found in regard to these putative epialleles. Instead, they covered diverse biological functions such as

DNA transcription, carbohydrate and lipid metabolic processes and photosynthesis (Table S35). Interestingly, both the gene body and its promoter were methylated in most putative genetic epialleles (Table S36). Most DMRs were less than 5 kbp away from a gene, suggesting a potential regulatory role (Fig. 7, Fig. S21, S22, S23, S24). Significant correlations using a 0.05  $p$  value threshold were detected when the influence of gene body methylation ( $\tau = -0.14$ ) and gene promoter methylation ( $\tau = -0.24$ ) on gene expression based on Kendall  $\tau$  tests. Both interactions show a negative correlation coefficient indicating that expression tended to decrease as methylation levels rose. The coefficients nonetheless do not imply a strictly strong relationship as expected due to the diversity of factors affecting gene expression and as also observed in the spectrum of gene expression and methylation interactions during seed formation (Fig. S25, S26, S27) and as also observed in other studies in tomatoes and strawberries (Lang et al. 2017; Cheng et al. 2018). However, during early seed development in

**Fig. 8** Gene expression and gene body and promoter methylation in CpG and CHG contexts from 15 days after pollination ovules displaying transgressive patterns in the  $F_0$  and its parents. Genes are sorted in the same order in both heatmaps

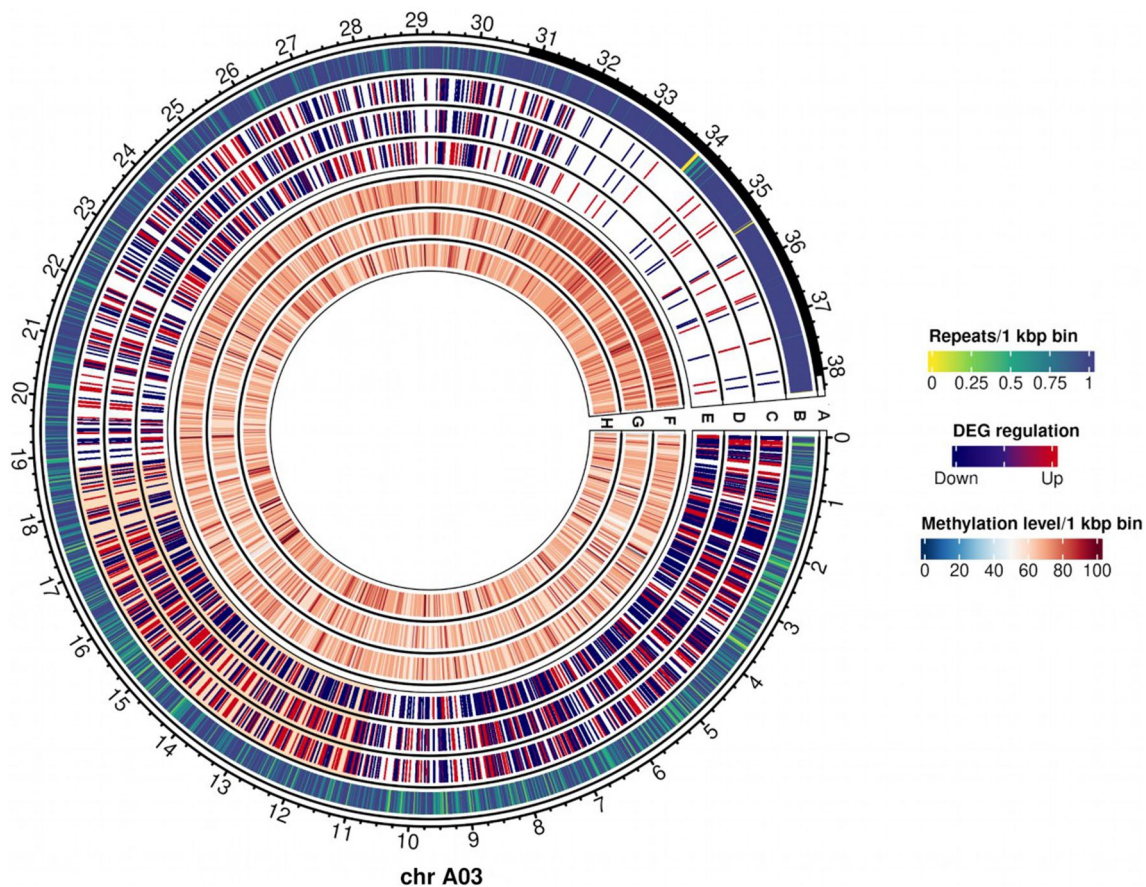


the hybrid, proportional interactions with upregulation of hypomethylated genes were most prevalent (Fig. 8). Studies linking methylation and expression and evaluating epialleles have proved beneficial in detecting heterotic patterns in other crops like maize, rice and Arabidopsis (Greaves et al. 2015; Cao et al. 2022; Wang and Wang 2022). In addition, 112,635 CpG islands were detected in all assigned chromosomes in the Express 617 reference genome, with an average length of 363 bp and strong differences in frequency in centromeric regions of different chromosomes (Table S37). Although 86% of all identified CpG islands were methylated, with an average methylation level of 62%, only 1.35% of these were differentially methylated (Table S38).

### Segmental and subgenome expression bias in hybrids

Segmental patterns of differential gene expression were visualized by circos plots displaying expression patterns for each chromosome, genotype and stage. The presence of putative expression clusters was assessed more precisely through a 500 kbp genome-wide binning approach where consistent DEGs patterns per segment, chromosome, genotype and stage were grouped as described in Material and Methods. Consequently, 144 differentially expressed segments across genotypes and stages were determined (Table S39). More differentially expressed segments were found in the A subgenome than the C subgenome and most segments found in  $F_0$  comparisons mimicked the expression patterns of the maternal parent Express 617 (Table S39). An example on chromosome A03 is shown in Fig. 9. Sequence reads from

G3D001 pollinated ovules were employed to exclude the possibility that the observed patterns were due to genomic rearrangements (Table S40). Large-scale deletions were found only in chromosome C01 in G3D001, which accounts for the low expression found on the deleted segments in that chromosome (Fig. S28, S29, S30, Table S39). However, no large-scale rearrangements were found in chromosome A03 in G3D001 (Fig. S31), so that cannot be the reason why a segment on this chromosome showed low expression in both early (15 days) and late (30 days) seed development stages (Fig. 9 and Fig. S32). Moreover, the corresponding chromosome region in Express 617 does not appear to be duplicated, since neither the  $F_1$  nor the  $F_0$  showed a high expression pattern that could have been inherited from a large-scale duplication from the Express 617 parent (Fig. S33, S34). Furthermore, no relationships were observed between methylation level, repeat density or position relative to the centromere. This suggests that specific chromosome segments may correspond more closely to maternal expression patterns than other regions. The mechanisms of such a phenomenon could be associated with parental roles during embryo development, genomic imprinting or chromatin activity and/or genome accessibility for transcription. However, more detailed investigations are necessary to validate these hypotheses.



**Fig. 9** Differentially expressed genes (DEGs) and methylation levels from 15 days after pollination ovules from  $F_0$  and parents in chromosome A03. Outer to inner tracks correspond to: (a) Predicted centromere positions in black; (b) Repeat density per 1 kbp bin; (c–e)

DEG regulation in (c) Express 617, (d)  $F_0$  and (e) G3D001; (SSf–h): Methylation levels per 1 kbp bin in (f) Express 617, (g)  $F_0$  and (h) G3D001. A differentially expressed chromosome segment between around 11 Mbp and 18.8 Mbp is highlighted in orange in tracks c–e

## Discussion

The results of this study demonstrate that differential expression and methylation patterns potentially associated with heterotic patterns are already detectable during seed development and seedling stages in an  $F_1$  hybrid. Most DEGs showed maternal and paternal dominances regardless of the tissue and stage, indicating that some of these differential features represent stable regulatory patterns with a general involvement in heterosis. Similar kinds of parental gene expression dominance have been reported previously in oilseed rape and cotton (Yoo et al. 2013; Wu et al. 2018; Wei et al. 2021). Li et al. (2020) demonstrated that expression dominance levels in interspecific hybrids can vary based on the sampled tissue, with stems and leaves showing more additive gene expression in allopolyploid *B. napus* compared to the expression from its diploid progenitors *B. rapa* and *B. oleracea*. Gene expression additivity was also reported by Zhang et al. (2021a) as a main pattern of expression dominance levels in excised

pod sections in crosses between *Raphanus sativus* (RR,  $2n=18$ ) and *B. oleracea*, whereas seeds and pods from the homozygous diploids displayed predominantly paternal dominance. The diversity of sampled tissues, species and genotypes in the previous studies and ours could account for the contrasting expression dominance levels observed. The segmental patterns of differential gene expression on a chromosome scale which we observed reflect similar results reported in *B. napus* by Lloyd et al. (2018) and He et al. (2017), who found that large-scale rearrangements induced large segmental expression differences. Interestingly, not all differentially expressed chromosome segments analyzed in our study were associated with large-scale rearrangements. Although the underlying reasons for this observation require further elucidation, this may indicate that chromosome-level patterns of chromatin rearrangement and transcription accessibility may be involved.

Additionally, pollinated  $F_0$  ovules that would develop later into  $F_1$  seeds and plants showed a higher similarity to the maternal genotype Express 617 in terms of gene

expression, small RNA expression and methylation. As noted by Jahnke et al. (2010), this might be due to the triploid nature of the endosperm, which arises from the union of a duplicated maternal gamete and a paternal gamete via double-fertilization. Seeds are composed of a seed coat, embryo and endosperm, with the proportions of the three components varying depending on species and developmental age. Transcriptomic profiling of different seed tissues using laser microdissection has been employed to characterize the transcriptomic profiles during seed formation in *A. thaliana* and *B. napus* (Kirkbride et al. 2019; Ziegler et al. 2019; Khan et al. 2022), and could provide further insights into heterotic patterns during  $F_1$  seed development; however, this technically demanding task was outside of the scope of the present study.

Subgenomic expression bias has been reported earlier in *Brassica* species (Bird et al. 2018, 2021a; Zhang et al. 2021b). Hence, we also investigated expression bias of differentially expressed up- and downregulated genes on a subgenomic basis in each genotype and stage. Although we did not detect any subgenome bias in gene expression, on the whole, more genes were differentially upregulated in G3D001 than in Express 617 (Table S16, Figures S11, S12, S13, S14, S15). The observed genotype-specific bias is potentially a result of different genomic, transcriptomic and epigenomic factors. Firstly, genomic rearrangements gene copy number variations and other structural variants are known to affect various traits in *B. napus* and other polyploid plants (Schiessl et al. 2017; Vollrath et al. 2021; Makhoul et al. 2022) and could have led to potential biases in expression patterns. Transcriptomic aspects such as gene isoforms, gene network interactions and allele expression bias might also be involved in favoring the up- or downregulation from a certain genotype or haplotype (Fan et al. 2020; Schiessl et al. 2020; Golicz et al. 2021). Lastly, epigenomic factors like parental gamete methylation mechanisms, genomic imprinting or differences between parental *cis-trans* regulating factors, miRNA isoforms (isomiRs) and TE families and densities could all result in potential genotype-biased or haplotype-biased expression (Jain et al. 2018; Go and Civetta 2020; Gill et al. 2021).

Around 12–18% of features shared in at least 2–3 stages followed maternal dominant patterns, highlighting the potential relevance of maternal effects on transcriptomic and epigenomic regulation of early development in this hybrid. Furthermore, less than 3% of all expression and methylation features had the same expression and methylation patterns across all stages, suggesting that the role of those features might be more essential throughout seed and early seedling development. Our results in regard to GO enrichment of differentially regulated features also underlined the role of heterotic patterns in driving key biological functions involved in photosynthesis, stress response, growth and

development. This also highlights the potential of RNA-Seq for global transcriptomic profiling as also supported by its high accuracy and robustness (Everaert et al. 2017; Coenye 2021). Furthermore, heterosis has been associated with a combination of similar functions like photosynthetic activity and cell division, which already help to enhance performance during early developmental stages (Liu et al. 2021). Here, DEGs involved in reproduction and meiotic functions already showed transgressive upregulated expression in ovules from selfed- $F_1$  plants at 15 days after pollination. Furthermore, the hybrids revealed a more robust structure and higher dry seed weight in comparison to their parents. Enrichment of GO groups linked to cell division, stress response and development functions, such as the ones detected in the hybrids of this study, have also led to similar phenotypes in hybrids of oilseed rape, cotton and Arabidopsis (Shen et al. 2017; Yang et al. 2017; Shahzad et al. 2020; Zhu et al. 2020; Rong et al. 2021). Information on these genes and their expression patterns could be of potential interest for proteomic validation and approaches using transcriptomic data for hybrid performance prediction.

Interestingly, we observed differentially expressed miRNAs in early and late seed development among miRNA families which are normally involved in plant growth and development (Plotnikova et al. 2019; Dong et al. 2022; Verma et al. 2022); thus, broadening the range and data availability of miRNAs in oilseed rape development, for example, miR172 regulates not only the flowering time pathway, but also embryo development by controlling *APETALA 2 (AP2)* and *AP2-like* genes such as *TOE2* (Boutillier et al. 2002; Shivaraj et al. 2018; Nowak et al. 2022) was found in our study. miR165/166 families control leaf adaxial/abaxial development and embryogenesis by targeting the class III homeodomain leucine zipper (HD-ZIP III) transcription factor gene family which includes the *REV*, *PHV* and *PHB* (Wang et al. 2007; Tang et al. 2012). Both *PHB* and *PHV* were identified in the present study and have been described to indirectly regulate *LEC2*, a gene that promotes embryo formation in Arabidopsis and seed size and seed lipid biosynthesis in *B. napus* (Braybrook et al. 2006; Tang et al. 2012; Wójcik et al. 2017; Miller et al. 2019). As noted by Dong et al. (2022), miR169 targets *C-REPEAT BINDING FACTOR (CBF)* and *NUCLEAR FACTOR YA (NF-YA)* genes. During late seed development in *B. napus*, we found that miR169 targets *EMB2016*, a member of the EMB gene family critical for embryo development (Tzafrir et al. 2004; Růžička et al. 2017; Meinke 2020), while *EMB2204* was targeted by miR3629. miR3629 was first reported in *Vitis vinifera* cv. Pinot Noir by Pantaleo et al. (2010) and has since been reported in *Camellia azalea*, in response to chilling in *Prunus persica* and in disease susceptibility in *V. vinifera* cv. Bosco and *V. vinifera* cv. Chardonnay (Barakat et al. 2012; Pantaleo et al. 2016; Yin

et al. 2016; Snyman et al. 2017). mir9410 has been detected in *B. oleracea* and *B. rapa* (Lukasik et al. 2013; Zhang et al. 2018), yet no clear function information for mir9410 exists for Brassica species. In our study, miR9410 targeted a *filamentation temperature sensitive protein H 1 (FtsH7)* gene copy encoding a protease that in turns degrades *D1* protein in photosystem II. *FtsH* genes have been reported in tomato, sorghum, Arabidopsis and *B. napus* (Xu et al. 2021; Yi et al. 2022).

The study identified multiple differential miRNA sequences and their putative targets with implications on plant development and performance. Further validation of targets associated with DE-miRNAs can potentially be achieved through degradome sequencing (German et al. 2008), precise isomiRs classification (Morin et al. 2008; Sablok et al. 2015; Yang et al. 2019), target knock-out experiments (Jain et al. 2018; Wei et al. 2018; Li et al. 2021) or gene co-expression networks (Schiessl et al. 2020) in order to delimit their role in seed and embryo formation in *B. napus*.

Overall, the number of methylated cytosines in the CHH context was higher in all genotypes compared to other contexts; nonetheless, methylation levels were higher in CpG and CHG contexts, as observed previously in multiple plants species (e.g. Niederhuth et al. 2016; Bartels et al. 2018). Methylation is generally associated with gene downregulation through transcription inhibition. Nevertheless, hypermethylation and hypomethylation were also linked with up- and downregulation, respectively. Proportional gene hypermethylation and upregulation were observed in mice and human cells (Arechederra et al. 2018; Rauluseviciute et al. 2020) as well as in strawberry and tomatoes (Lang et al. 2017; Cheng et al. 2018); however, no mechanisms explaining gene activation through hypermethylation are fully known so far; thus, further research would elucidate the interactions between methylation and gene regulation, particularly in relation to heterozygosity and heterosis.

In our study, we evaluated methylation during seed development because parental asymmetric methylation and genomic imprinting occurs mostly at that stage in flowering plants (Batista and Köhler 2020). DNA hypomethylation of the female gamete and paternal gamete hypermethylation has been reported in many flowering plants, including Arabidopsis, rice and maize (Gehring et al. 2009; Zemach et al. 2010; Zhang et al. 2014b). Interestingly, we observed contrasting patterns of maternal hypermethylation and paternal hypomethylation in the  $F_0$ . Similar parental methylation trends were also observed in the  $F_1$  despite allele segregation. Such patterns were also reported by Liu et al. (2018) in *B. napus* and by Grover et al. (2020) in *B. rapa*. As a possible explanation for maternal hypermethylation, Grover et al. (2020) proposed that a high expression of

so-called siren siRNAs in the seed coat could trigger maternal DNA methylation during seed development. The molecular mechanisms and effects of genomic imprinting, where an allele follows a parental expression pattern due to inherited epigenomic modifications, are restricted mostly to the endosperm rather than the embryo in flowering plants and have been extensively discussed by Weigel and Colot (2012) and Batista and Köhler (2020), respectively. The role of imprinted genes has been linked to chromatin modification, hormone biosynthesis, nutrient transfer, endosperm proliferation and seed size regulation (reviewed by Jiang and Köhler 2012; Batista and Köhler 2020). Furthermore, Rong et al. (2021) reported the enrichment of transposable elements located in 5 kbp flanking regions of imprinted genes in *B. napus*. Cao et al. (2022) analyzed imprinted genes in six backcrossing generations of maize as well as in three selfing generations derived from the 6th backcross. They proposed that the divergence between TEs derived from 24-nt siRNAs in the parental maize genomes might have led to transgenerational inheritance of imprinted genes. Putative imprinted genes were also found in the seedling and seed development stages in our study. Epigenetic changes have been reported as relevant heterotic factors which are influenced by allele diversity, parental effects and environmental conditions (Botet and Keurentjes 2020). Epigenomic parental effects are more likely to occur during seed formation, when gametes fuse to form a zygote, given that this stage is marked by epigenomic features involving siRNA, DNA methylation, imprinting and chromatin activity. Because these features shape the epigenomic and transcriptomic landscape of the zygote and, potentially, its future development as a seedling, hybrids can potentially benefit strongly from a heterotic advantage imparted by these features in these very early developmental stages.

Most frequently, differentially methylated transposable elements corresponded to abundant Copia and Gypsy families. Transposable elements are key factors in speciation and subgenome expression patterns (Bird et al. 2018, 2021a; Bottani et al. 2018) and are known for their high variability across plant species (Novák et al. 2020; Mhiri et al. 2022). TEs can also alter the epigenetic landscape in relation to hybrid fitness (Serrato-Capuchina and Matute 2018). Therefore, detailed assessment of transposable element densities and compositions between hybrids and their parents could be beneficial. In addition, siRNAs are known to mediate silencing of transposable elements via the RNA-directed DNA methylation (RdDM) pathway. At the same time, TEs are a source of sRNAs, including siRNAs, that could potentially silence TEs through a post-transcriptional gene silencing (PTGS) process (Matzke and Mosher 2014; Gill et al. 2021). As also described by Rong et al. (2021), most differentially methylated TEs were found

in 5 kbp gene flanking regions. Because most TEs, DMRs and DE-siRNAs converged into regions directly flanking genes, whereas there was no abundance in regions flanking DEGs, gene regulation in the hybrid appears to be associated with conservation of key genetic functions by (i) maintaining similar gene expression patterns between genotypes despite differential methylation in gene flanking regions for most genes and (ii) reducing the number of DMRs, DE-siRNAs and differentially methylated TEs in the proximity of DEGs similarly to other reports in poplar, mammals and invertebrates (Blake et al. 2020; Zhang et al. 2020; Cardoso-Júnior et al. 2021; Dixon and Matz 2022). This could be one of the most interesting and significant results of the present study. Methylation data have been employed in plant breeding to investigate and improve flowering time, disease susceptibility and abiotic stress response (Mercé et al. 2020; Shaikh et al. 2022), and have enabled genomic diversity expansion and discovery of cis-regulatory elements (Xu et al. 2019; Crisp et al. 2020). The present study highlights the spectrum of methylation patterns and putative epialleles in oilseed rape with special implications on seed formation. Additionally, most CpG islands were not differentially methylated between genotypes, indicating a putative role in regulating gene expression, as reported in rice and *Arabidopsis* (Ashikawa 2001).

The sequencing data gathered from single tissues in the present study allowed an integrated view of gene expression, small RNA interactions and genomic methylation during early developmental stages in an oilseed rape hybrid developed from two distant genotypes. Coding and non-coding features which were differentially expressed or methylated in this study provide new insights into early expression of heterosis in oilseed rape seeds and seedlings from a molecular viewpoint and constitute an extensive multiomics atlas for oilseed rape breeding. The extent of these features in an allopolyploid model crop like *B. napus* also have potential implications in other polyploid crops where heterosis still remains to be exploited, such as wheat and potatoes (Steeg et al. 2022). Patterns of expression and methylation dominance levels could also contribute a new level of understand regarding allele-specific gene expression (Fan et al. 2020; Sands et al. 2021), isoform expression (Vitting-Seerup and Sandelin 2019; Yao et al. 2020; Golicz et al. 2021), gene fusion and dosage (Mahmoud et al. 2019; Serin Harmanci et al. 2020; Bird et al. 2021b) as well as non-germline omics variations among  $F_1$  plants and populations (Higgins et al. 2018; Cortijo et al. 2019; Orantes-Bonilla et al. 2022; Quezada-Martinez et al. 2022). Their role in heterotic gene expression patterns is ultimately also of interest for transcriptome-based genomic selection or hybrid performance prediction (e.g. Frisch et al. 2010). Defining the roles of differentially expression regulatory features in early developmental stages of hybrids could be

used to enhance expression-based prediction model (Seifert et al. 2018b; Zrimec et al. 2020; Cheng et al. 2021; Hu et al. 2021c; Knoch et al. 2021). Altogether our findings highlight transcriptomics and epigenomic differences between early developmental stages in  $F_1$  and  $F_0$  in terms of methylation level as well as in gene and small RNA expression. The contribution of differential coding and non-coding features to early hybrid seed formation is of key interest for hybrid breeding and deserves further evaluation using more diverse genotypes, heterotic pools and species. Future developments in sequencing and bioinformatics will also aid in elucidating the role and interactions among transcriptomic and epigenomic features at higher resolution, helping to expand current knowledge and applications of heterosis in polyploid crops.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00122-023-04345-7>.

**Acknowledgements** The authors acknowledge support for bioinformatic resources from the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) and the Bioinformatics Core Facility at JLU. The experimental design was created with BioRender.com.s

**Author contributions statement** R.J.S. and J.Z. conceived and supervised the study. M.O.B. drafted the manuscript and designed the bioinformatic analyses. H.W. conducted bioinformatic studies, generated crosses, extracted and sampled pollinated ovules for sequencing and contributed to data analysis. D.H. carried and supervised the field experiment. W.L. contributed to experimental trials. H.T.L. performed genome synteny analyses and contributed to whole-genome bisulfite and mRNA analyses. A.A.G. contributed to analyses of transcriptomic and epigenomic features. All authors read and approved the manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This work was performed within the framework of the Joint Sino-German Research Program (2018) with support from grant SN14/22-1 to RJS from the German Research Foundation (DFG) and grant 31861133016 to JZ from the National Natural Science Foundation of China (NSFC).

**Data availability** The mRNA, sRNA and WGBS libraries and fragment count datasets generated in this study are found in the GEO data repository under accession GSE202610. G3D001 genomic reads from self-pollinated ovules are found in NCBI Bioproject PRJNA850551.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended

use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alonso-Peral MM, Trigueros M, Sherman B, Ying H, Taylor JM, Peacock WJ et al (2017) Patterns of gene expression in developing embryos of *Arabidopsis* hybrids. *Plant J* 89:927–939. <https://doi.org/10.1111/tpj.13432>
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Accessed 2021
- Arechederra M, Daian F, Yim A, Bazai SK, Richelme S, Dono R et al (2018) Hypermethylation of gene body CpG islands predicts high dosage of functional oncogenes in liver cancer. *Nat Commun* 9:3164. <https://doi.org/10.1038/s41467-018-05550-5>
- Ashikawa I (2001) Gene-associated CpG islands in plants as revealed by analyses of genomic sequences. *Plant J* 26:617–625. <https://doi.org/10.1046/j.1365-313x.2001.01062.x>
- Axtell MJ (2013) Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol* 64:137–159. <https://doi.org/10.1146/annurev-arplant-050312-120043>
- Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A et al (2016) Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* 44:W147–W153. <https://doi.org/10.1093/nar/gkw419>
- Bao J, Lee S, Chen C, Zhang X, Zhang Y, Liu S et al (2005) Serial analysis of gene expression study of a hybrid rice strain (LYP9) and its parental cultivars. *Plant Physiol* 138:1216–1231. <https://doi.org/10.1104/pp.105.060988>
- Barakat A, Sriram A, Park J, Zhebentyayeva T, Main D, Abbott A (2012) Genome wide identification of chilling responsive microRNAs in *Prunus persica*. *BMC Genomics* 13:481. <https://doi.org/10.1186/1471-2164-13-481>
- Bartels A, Han Q, Nair P, Stacey L, Gaynier H, Mosley M et al (2018) Dynamic DNA methylation in plant growth and development. *Int J Mol Sci* 19:2144. <https://doi.org/10.3390/ijms19072144>
- Basunanda P, Radoev M, Ecke W, Friedt W, Becker HC, Snowdon RJ (2010) Comparative mapping of quantitative trait loci involved in heterosis for seedling and yield traits in oilseed rape (*Brassica napus* L.). *Theor Appl Genet* 120:271–281. <https://doi.org/10.1007/s00122-009-1133-z>
- Batista RA, Köhler C (2020) Genomic imprinting in plants-revisiting existing models. *Genes Dev* 34:24–36. <https://doi.org/10.1101/gad.332924.119>
- Bird KA, VanBuren R, Puzey JR, Edger PP (2018) The causes and consequences of subgenome dominance in hybrids and recent polyploids. *New Phytol* 220:87–93. <https://doi.org/10.1111/nph.15256>
- Bird KA, Niederhuth CE, Ou S, Gehan M, Pires JC, Xiong Z et al (2021a) Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid *Brassica napus*. *New Phytol* 230:354–371. <https://doi.org/10.1111/nph.17137>
- Bird KA, Pires JC, VanBuren R, Xiong Z, Edger PP (2021b) Gene balance in allopolyploids: homoeologous exchanges show signs of dosage constraint and dosage constraint of biased homoeologs differs between subgenomes. *BioRxiv*. <https://doi.org/10.1101/2021.11.16.468838>
- Blake LE, Roux J, Hernando-Herraez I, Banovich NE, Perez RG, Hsiao CJ et al (2020) A comparison of gene expression and DNA methylation patterns across tissues and species. *Genome Res* 30:250–262. <https://doi.org/10.1101/gr.254904.119>
- Bonjean AP, Dequidt C, Sang T (2016) Rapeseed in China. *OCL* 23:D605. <https://doi.org/10.1051/ocl/2016045>
- Botet R, Keurentjes JJB (2020) The role of transcriptional regulation in hybrid vigor. *Front Plant Sci* 11:410. <https://doi.org/10.3389/fpls.2020.00410>
- Bottani S, Zabet NR, Wendel JF, Veitia RA (2018) Gene Expression dominance in Allopolyploids: hypotheses and models. *Trends Plant Sci* 23:393–402. <https://doi.org/10.1016/j.tplants.2018.01.002>
- Boutillier K, Offringa R, Sharma VK, Kieft H, Ouellet T, Zhang L et al (2002) Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14:1737–1749. <https://doi.org/10.1105/tpc.001941>
- Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL et al (2006) Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc Natl Acad Sci U S A* 103:3468–3473. <https://doi.org/10.1073/pnas.0511331103>
- BSA (2022) Federal Plant Variety Office (BSA) descriptive variety 2022 list: cereals, maize, oil and fibre plants, legumes, beets, catch crops. [https://www.bundessortenamt.de/bsa/media/Files/BSL/bsl\\_getreide\\_2022.pdf](https://www.bundessortenamt.de/bsa/media/Files/BSL/bsl_getreide_2022.pdf). Accessed 2022
- Cao S, Wang L, Han T, Ye W, Liu Y, Sun Y et al (2022) Small RNAs mediate transgenerational inheritance of genome-wide transacting epialleles in maize. *Genome Biol* 23:53. <https://doi.org/10.1186/s13059-022-02614-0>
- Cardoso-Júnior CA, Yagound B, Ronai I, Remnant EJ, Hartfelder K, Oldroyd BP (2021) DNA methylation is not a driver of gene expression reprogramming in young honey bee workers. *Mol Ecol* 30:4804–4818. <https://doi.org/10.1111/mec.16098>
- Castillo-Bravo R, Fort A, Cashell R, Brychkova G, McKeown PC, Spillane C (2022) Parent-of-origin effects on seed size modify heterosis responses in *Arabidopsis thaliana*. *Front Plant Sci* 13:835219. <https://doi.org/10.3389/fpls.2022.835219>
- Catoni M, Tsang JM, Greco AP, Zabet NR (2018) DMRcaller: a versatile R/bioconductor package for detection and visualization of differentially methylated regions in CpG and non-CpG contexts. *Nucleic Acids Res* 46:e114. <https://doi.org/10.1093/nar/gky602>
- Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X et al (2014) Plant genetics. Early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. *Science* 345:950–953. <https://doi.org/10.1126/science.1253435>
- Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Chen L, Zhu Y, Ren X, Yao D, Song Y, Fan S et al (2022) Heterosis and differential DNA methylation in soybean hybrids and their parental lines. *Plants Basel* 11:1136. <https://doi.org/10.3390/plants11091136>
- Cheng C-Y, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD (2017) Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J* 89:789–804. <https://doi.org/10.1111/tpj.13415>
- Cheng J, Niu Q, Zhang B, Chen K, Yang R, Zhu JK et al (2018) Down-regulation of RdDM during strawberry fruit ripening. *Genome Biol* 19:1–14. <https://doi.org/10.1186/s13059-018-1587-x>
- Cheng C-Y, Li Y, Varala K, Bubert J, Huang J, Kim GJ et al (2021) Evolutionarily informed machine learning enhances the power

- of predictive gene-to-phenotype relationships. *Nat Commun* 12:5627. <https://doi.org/10.1038/s41467-021-25893-w>
- Church BV, Williams HT, Mar JC (2019) Investigating skewness to understand gene expression heterogeneity in large patient cohorts. *BMC Bioinform* 20:1–14. <https://doi.org/10.1186/s12859-019-3252-0>
- Coenye T (2021) Do results obtained with RNA-sequencing require independent verification? *Biofilm* 3:100043. <https://doi.org/10.1016/j.biofilm.2021.100043>
- Cortijo S, Aydin Z, Ahnert S, Locke JC (2019) Widespread inter-individual gene expression variability in *Arabidopsis thaliana*. *Mol Syst Biol* 15:e8591
- Crisp PA, Marand AP, Noshay JM, Zhou P, Lu Z, Schmitz RJ et al (2020) Stable unmethylated DNA demarcates expressed genes and their cis-regulatory space in plant genomes. *Proc Natl Acad Sci* 117:23991–24000. <https://doi.org/10.1073/pnas.2010250117>
- Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM et al (2022) Ensembl 2022. *Nucleic Acids Res* 50:D988–D995. <https://doi.org/10.1093/nar/gkab1049>
- Dai X, Zhuang Z, Zhao PX (2018) psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res* 46:W49–W54. <https://doi.org/10.1093/nar/gky316>
- Dainat J (2019) *AGAT: Another Gff analysis toolkit to handle annotations in any GTF/GFF format. (Version v0.5.0)*. 10.5281/zenodo.3552717. Accessed 2020
- Darwin C (1876) The effects of cross and self fertilisation in the vegetable kingdom. John Murry, London
- Di Y, Schafer DW, Cumbie JS, Chang JH (2011) The NBP negative binomial model for assessing differential gene expression from RNA-Seq. *Stat Appl Genet Mol Biol* 10:1–28. <https://doi.org/10.2202/1544-6115.1637>
- Dixon G, Matz M (2022) Changes in gene body methylation do not correlate with changes in gene expression in Anthozoa or Hexapoda. *BMC Genomics* 23:1–11. <https://doi.org/10.1186/s12864-022-08474-z>
- Dong Q, Hu B, Zhang C (2022) microRNAs and their roles in plant development. *Front Plant Sci* 13:824240. <https://doi.org/10.3389/fpls.2022.824240>
- Doyle JJ, Doyle J (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15
- Everaert C, Luypaert M, Maag JL, Cheng QX, Dinger ME, Hellemans J et al (2017) Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. *Sci Rep* 7:1–11. <https://doi.org/10.1038/s41598-017-01617-3>
- Ewels P, Magnusson M, Lundin S, Källner M (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32:3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>
- Fan J, Hu J, Xue C, Zhang H, Susztak K, Reilly MP et al (2020) ASEP: Gene-based detection of allele-specific expression across individuals in a population by RNA sequencing. *PLoS Genet* 16:e1008786. <https://doi.org/10.1371/journal.pgen.1008786>
- FAO (2022). FAOSTAT. <https://www.fao.org/faostat/en>. Accessed 2022.
- Fitz-James MH, Cavalli G (2022) Molecular mechanisms of transgenerational epigenetic inheritance. *Nat Rev Genet* 23:325–341. <https://doi.org/10.1038/s41576-021-00438-5>
- Friedt W, Tu J, Fu T (2018) Academic and economic importance of *Brassica napus* rapeseed. In: Liu S, Snowdon R (eds) *The Brassica napus genome* chalhoub. Springer International Publishing, Cham, pp 1–20
- Frisch M, Thiemann A, Fu J, Schrag TA, Scholten S, Melchinger AE (2010) Transcriptome-based distance measures for grouping of germplasm and prediction of hybrid performance in maize. *Theor Appl Genet* 120:441–450. <https://doi.org/10.1007/s00122-009-1204-1>
- Fujimoto R, Uezono K, Ishikura S, Osabe K, Peacock WJ, Dennis ES (2018) Recent research on the mechanism of heterosis is important for crop and vegetable breeding systems. *Breed Sci* 68:145–158. <https://doi.org/10.1270/jsbbs.17155>
- Gardiner-Garden M, Frommer M (1987) CpG Islands in vertebrate genomes. *J Mol Biol* 196:261–282. [https://doi.org/10.1016/0022-2836\(87\)90689-9](https://doi.org/10.1016/0022-2836(87)90689-9)
- Ge SX, Son EW, Yao R (2018) iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinform* 19:534. <https://doi.org/10.1186/s12859-018-2486-6>
- Ge SX, Jung D, Yao R (2020) ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36:2628–2629. <https://doi.org/10.1093/bioinformatics/btz931>
- Gehring M, Bubb KL, Henikoff S (2009) Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* 324:1447–1451. <https://doi.org/10.1126/science.1171609>
- German MA, Pillay M, Jeong D-H, Hetawal A, Luo S, Janardhanan P et al (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol* 26:941–946. <https://doi.org/10.1038/nbt1417>
- Gill RA, Scossa F, King GJ, Golicz AA, Tong C, Snowdon RJ et al (2021) On the role of transposable elements in the regulation of gene expression and subgenomic interactions in crop genomes. *Crit Rev Plant Sci* 40:157–189. <https://doi.org/10.1080/07352689.2021.1920731>
- Gerke A, Schierholt A, Becker HC (2012) Extending the rapeseed gene pool with resynthesized *Brassica napus* II: Heterosis. *Theor Appl Genet* 124:1017–1026. <https://doi.org/10.1007/s00122-011-1765-7>
- Go AC, Civetta A (2020) Hybrid incompatibilities and transgressive gene expression between two closely related subspecies of *Drosophila*. *Front Genet* 11:599292. <https://doi.org/10.3389/fgene.2020.599292>
- Golicz AA, Allu AD, Li W, Lohani N, Singh MB, Bhalla PL (2021) A dynamic intron retention program regulates the expression of several hundred genes during pollen meiosis. *Plant Reprod* 34:225–242. <https://doi.org/10.1007/s00497-021-00411-6>
- Greaves IK, Gonzalez-Bayon R, Wang L, Zhu A, Liu P-C, Groszmann M et al (2015) Epigenetic changes in hybrids. *Plant Physiol* 168:1197–1205. <https://doi.org/10.1104/pp.15.00231>
- Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 34:D140–D144. <https://doi.org/10.1093/nar/gkj112>
- Groszmann M, Gonzalez-Bayon R, Greaves IK, Wang L, Huen AK, Peacock WJ et al (2014) Intraspecific *Arabidopsis* hybrids show different patterns of heterosis despite the close relatedness of the parental genomes. *Plant Physiol* 166:265–280. <https://doi.org/10.1104/pp.114.243998>
- Grover JW, Burgess D, Kendall T, Baten A, Pokhrel S, King GJ et al (2020) Abundant expression of maternal siRNAs is a conserved feature of seed development. *Proc Natl Acad Sci U S A* 117:15305–15315. <https://doi.org/10.1073/pnas.2001332117>
- Gu Z, Gu L, Eils R, Schlesner M, Brors B (2014) circlize Implements and enhances circular visualization in R. *Bioinformatics* 30:2811–2812. <https://doi.org/10.1093/bioinformatics/btu393>
- Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32:2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>
- He Z, Wang L, Harper AL, Havlickova L, Pradhan AK, Parkin IAP et al (2017) Extensive homoeologous genome exchanges in allopolyploid crops revealed by mRNAseq-based visualization. *Plant Biotechnol J* 15:594–604. <https://doi.org/10.1111/pbi.12657>

- Higgins EE, Clarke WE, Howell EC, Armstrong SJ, Parkin IAP (2018) Detecting de novo homoeologous recombination events in cultivated *Brassica napus* using a genome-wide SNP array. *G Bethesda* 8:2673–2683. <https://doi.org/10.1534/g3.118.200118>
- Hochholdinger F, Hoecker N (2007) Towards the molecular basis of heterosis. *Trends Plant Sci* 12:427–432. <https://doi.org/10.1016/j.tplants.2007.08.005>
- Hu D, Jing J, Snowdon RJ, Mason AS, Shen J, Meng J et al (2021a) Exploring the gene pool of *Brassica napus* by genomics-based approaches. *Plant Biotechnol J* 19:1693–1712. <https://doi.org/10.1111/pbi.13636>
- Hu D, Zhao Y, Shen J, He X, Zhang Y, Jiang Y et al (2021b) Genome-wide prediction for hybrids between parents with distinguished difference on exotic introgressions in *Brassica napus*. *Crop J* 9:1169–1178. <https://doi.org/10.1016/j.cj.2020.11.002>
- Hu H, Campbell MT, Yeats TH, Zheng X, Runcie DE, Covarrubias-Pazarán G et al (2021c) Multi-omics prediction of oat agronomic and seed nutritional traits across environments and in distantly related populations. *Theor Appl Genet* 134:4043–4054. <https://doi.org/10.1007/s00122-021-03946-4>
- Jahnke S, Sarholz B, Thiemann A, Kühn V, Gutiérrez-Marcos JF, Geiger HH et al (2010) Heterosis in early seed development: a comparative study of F1 embryo and endosperm tissues 6 days after fertilization. *Theor Appl Genet* 120:389–400. <https://doi.org/10.1007/s00122-009-1207-y>
- Jain A, Anand S, Singh NK, Das S (2018) Sequence and functional characterization of MIRNA164 promoters from Brassica shows copy number dependent regulatory diversification among homeologs. *Funct Integr Genomics* 18:369–383. <https://doi.org/10.1007/s10142-018-0598-8>
- Jian H, Yang B, Zhang A, Ma J, Ding Y, Chen Z et al (2018) Genome-wide identification of microRNAs in response to cadmium stress in oilseed rape (*Brassica napus* L.) using high-throughput sequencing. *Int J Mol Sci* 19:1431. <https://doi.org/10.3390/ijms19051431>
- Jian H, Zhang A, Ma J, Wang T, Yang B, Shuang LS et al (2019) Joint QTL mapping and transcriptome sequencing analysis reveal candidate flowering time genes in *Brassica napus* L. *BMC Genomics* 20:21. <https://doi.org/10.1186/s12864-018-5356-8>
- Jiang H, Köhler C (2012) Evolution, function, and regulation of genomic imprinting in plant seed development. *J Exp Bot* 63:4713–4722. <https://doi.org/10.1093/jxb/ers145>
- Johnson NR, Yeoh JM, Coruh C, Axtell MJ (2016) Improved placement of multi-mapping small RNAs. *G3 Bethesda* 6:2103–2111. <https://doi.org/10.1534/g3.116.030452>
- Kawanabe T, Ishikura S, Miyaji N, Sasaki T, Wu LM, Itabashi E et al (2016) Role of DNA methylation in hybrid vigor in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 113:E6704–E6711. <https://doi.org/10.1073/pnas.1613372113>
- Khan D, Ziegler DJ, Kalichuk JL, Hoi V, Huynh N, Hajihassani A et al (2022) Gene expression profiling reveals transcription factor networks and subgenome bias during *Brassica napus* seed development. *Plant J* 109:477–489. <https://doi.org/10.1111/tpj.15587>
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37:907–915. <https://doi.org/10.1038/s41587-019-0201-4>
- Kirkbride RC, Lu J, Zhang C, Mosher RA, Baulcombe DC, Chen ZJ (2019) Maternal small RNAs mediate spatial-temporal regulation of gene expression, imprinting, and seed development in *Arabidopsis*. *Proc Natl Acad Sci U S A* 116:2761–2766. <https://doi.org/10.1073/pnas.1807621116>
- Knoch D, Werner CR, Meyer RC, Riewe D, Abbadia A, Lücke S et al (2021) Multi-omics-based prediction of hybrid performance in canola. *Theor Appl Genet* 134:1147–1165. <https://doi.org/10.1007/s00122-020-03759-x>
- Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27:1571–1572. <https://doi.org/10.1093/bioinformatics/btr167>
- Krueger F, James F, Ewels P, Afyounian E and Schuster-Boeckler B (2021) *FelixKrueger/TrimGalore: v0.6.7*. *Zenodo*. 10.5281/zenodo.5127899
- Lai Z, Gross BL, Zou Y, Andrews J, Rieseberg LH (2006) Microarray analysis reveals differential gene expression in hybrid sunflower species. *Mol Ecol* 15:1213–1227. <https://doi.org/10.1111/j.1365-294X.2006.02775.x>
- Lang Z, Wang Y, Tang K, Tang D, Datsenko T et al (2017) 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* 114:E4511–E4519. <https://doi.org/10.1073/pnas.1705233114>
- Lauss K, Wardenaar R, Oka R, van Hulten MHA, Guryev V, Keurentjes JJB et al (2018) Parental DNA methylation states are associated with heterosis in epigenetic hybrids. *Plant Physiol* 176:1627–1645. <https://doi.org/10.1104/pp.17.01054>
- Lee H, Chawla HS, Obermeier C, Dreyer F, Abbadia A, Snowdon R (2020) Chromosome-scale assembly of winter oilseed rape *Brassica napus*. *Front Plant Sci* 11:496. <https://doi.org/10.3389/fpls.2020.00496>
- Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li A, Liu D, Wu J, Zhao X, Hao M, Geng S et al (2014) mRNA and small RNA transcriptomes reveal insights into dynamic homeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell* 26:1878–1900. <https://doi.org/10.1105/tpc.114.124388>
- Li J, Huang Q, Sun M, Zhang T, Li H, Chen B et al (2016) Global DNA methylation variations after short-term heat shock treatment in cultured microspores of *Brassica napus* cv. Topas Sic Rep 6:38401. <https://doi.org/10.1038/srep38401>
- Li M, Wang R, Wu X, Wang J (2020) Homeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid *Brassica napus*. *BMC Genomics* 21:330. <https://doi.org/10.1186/s12864-020-6747-1>
- Li Q, Shah N, Zhou X, Wang H, Yu W, Luo J et al (2021) Identification of micro ribonucleic acids and their targets in response to plasmodiophora brassicae infection in *Brassica napus*. *Front Plant Sci* 12:734419. <https://doi.org/10.3389/fpls.2021.734419>
- Li H (2016) *Seqtk*. <https://github.com/lh3/seqtk>. Accessed 2019
- Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30:923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Liu J, Li J, Liu H-F, Fan S-H, Singh S, Zhou X-R et al (2018) Genome-wide screening and analysis of imprinted genes in rapeseed (*Brassica napus* L.) endosperm. *DNA Res* 25:629–640. <https://doi.org/10.1093/dnares/dsy030>
- Liu W, He G, Deng XW (2021) Biological pathway expression complementation contributes to biomass heterosis in *Arabidopsis*. *Proc Natl Acad Sci U S A* 118:e2023278118. <https://doi.org/10.1073/pnas.2023278118>
- Lloyd A, Blary A, Charif D, Charpentier C, Tran J, Balzergue S et al (2018) Homeologous exchanges cause extensive dosage-dependent gene expression changes in an allopolyploid crop. *New Phytol* 217:367–377. <https://doi.org/10.1111/nph.14836>

- Louwaars NP (2018) Plant breeding and diversity: a troubled relationship? *Euphytica* 214:114. <https://doi.org/10.1007/s10681-018-2192-5>
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>
- Lukasik A, Pietrykowska H, Paczek L, Szweykowska-Kulinska Z, Zielenkiewicz P (2013) High-throughput sequencing identification of novel and conserved miRNAs in the *Brassica oleracea* leaves. *BMC Genomics* 14:801. <https://doi.org/10.1186/1471-2164-14-801>
- Lunardon A, Johnson NR, Hagerott E, Phifer T, Polydore S, Coruh C et al (2020) Integrated annotations and analyses of small RNA-producing loci from 47 diverse plants. *Genome Res* 30:497–513. <https://doi.org/10.1101/gr.256750.119>
- Ma J, Zhang D, Cao Y, Wang L, Li J, Lübberstedt T et al (2018) Heterosis-related genes under different planting densities in maize. *J Exp Bot* 69:5077–5087. <https://doi.org/10.1093/jxb/ery282>
- Mackay II, Cockram J, Howell P, Powell W (2021) Understanding the classics: the unifying concepts of transgressive segregation, inbreeding depression and heterosis and their central relevance for crop breeding. *Plant Biotechnol J* 19:26–34. <https://doi.org/10.1111/pbi.13481>
- Mahmoud M, Gobet N, Cruz-Dávalos DI, Mounier N, Dessimoz C, Sedlazeck FJ (2019) Structural variant calling: the long and the short of it. *Genome Biol* 20:246. <https://doi.org/10.1186/s13059-019-1828-7>
- Makhoul M, Chawla HS, Wittkop B, Stahl A, Voss-Fels KP, Zetzsche H et al (2022) Long-amplicon single-molecule sequencing reveals novel, trait-associated variants of vernalization1 homologs in hexaploid wheat. *Front Plant Sci* 13:942461. <https://doi.org/10.3389/fpls.2022.942461>
- Malla S, Brewin DG (2019) Crop research, biotech canola, and innovation policy in Canada: challenges, opportunities, and evolution. *Can J Agric Econ Revue Can D'agroéconomie* 67:135–150. <https://doi.org/10.1111/cjag.12195>
- Martinez Palacios P, Jacquemot M-P, Tapie M, Rousselet A, Diop M, Remoué C et al (2019) Assessing the response of small RNA populations to Allopolyploidy using resynthesized *Brassica napus* Allotetraploids. *Mol Biol Evol* 36:709–726. <https://doi.org/10.1093/molbev/msz007>
- Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* 15:394–408. <https://doi.org/10.1038/nrg3683>
- Meinke DW (2020) Genome-wide identification of EMBRYO-DEFECTIVE (EMB) genes required for growth and development in *Arabidopsis*. *New Phytol* 226:306–325. <https://doi.org/10.1111/nph.16071>
- Mercé C, Bayer PE, Tay Fernandez C, Batley J, Edwards D (2020) Induced methylation in plants as a crop improvement tool: progress and perspectives. *Agronomy* 10:1484. <https://doi.org/10.3390/agronomy10101484>
- Meyer RC, Witucka-Wall H, Becher M, Blacha A, Boudichevskaia A, Dörmann P et al (2012) Heterosis manifestation during early *Arabidopsis* seedling development is characterized by intermediate gene expression and enhanced metabolic activity in the hybrids. *Plant J* 71:669–683. <https://doi.org/10.1111/j.1365-313X.2012.05021.x>
- Mhiri C, Borges F, Grandbastien M-A (2022) Specificities and dynamics of transposable elements in land plants. *Biol Basel* 11:448. <https://doi.org/10.3390/biology11040488>
- Miller C, Wells R, McKenzie N, Trick M, Ball J, Fathi A et al (2019) Variation in expression of the HECT E3 ligase UPL3 modulates LEC2 levels, seed size, and crop yields in *Brassica napus*. *Plant Cell* 31:2370–2385. <https://doi.org/10.1105/tpc.18.00577>
- Montgomery SA, Berger F (2021) The evolution of imprinting in plants: beyond the seed. *Plant Reprod* 34:373–383. <https://doi.org/10.1007/s00497-021-00410-7>
- Morin RD, Aksay G, Dolgosheina E, Ebhardt HA, Magrini V, Mardis ER et al (2008) Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Res* 18:571–584. <https://doi.org/10.1101/gr.6897308>
- Muyle AM, Seymour DK, Lv Y, Huettel B, Gaut BS (2022) Gene body methylation in plants: mechanisms, functions, and important implications for understanding evolutionary processes. *Genome Biol Evol* 14:evac038. <https://doi.org/10.1093/gbe/evac038>
- Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q et al (2016) Widespread natural variation of DNA methylation within angiosperms. *Genome Biol* 17:194. <https://doi.org/10.1186/s13059-016-1059-0>
- Novák P, Guignard MS, Neumann P, Kelly LJ, Mlinarec J, Koblížková A et al (2020) Repeat-sequence turnover shifts fundamentally in species with large genomes. *Nat Plants* 6:1325–1329. <https://doi.org/10.1038/s41477-020-00785-x>
- Nowak K, Morończyk J, Grzyb M, Szczygieł-Sommer A, Gaj MD (2022) miR172 regulates WUS during somatic embryogenesis in *Arabidopsis* via AP2. *Cells* 11:718. <https://doi.org/10.3390/cells11040718>
- O'Brien KP, Remm M, Sonnhammer ELL (2005) Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res* 33:D476–D480. <https://doi.org/10.1093/nar/gki107>
- Orantes-Bonilla M, Makhoul M, Lee H, Chawla HS, Vollrath P, Langstroff A et al (2022) Frequent spontaneous structural rearrangements promote rapid genome diversification in a *Brassica napus* F1 generation. *Front Plant Sci* 13:1057953. <https://doi.org/10.3389/fpls.2022.1057953>
- Pantaleo V, Szittyá G, Moxon S, Miozzi L, Moulton V, Dalmay T et al (2010) Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. *Plant J No No*. <https://doi.org/10.1111/j.1365-313X.2010.04208.x>
- Pantaleo V, Vitali M, Boccacci P, Miozzi L, Cuozzo D, Chitarra W et al (2016) Novel functional microRNAs from virus-free and infected *Vitis vinifera* plants under water stress. *Sci Rep* 6:20167. <https://doi.org/10.1038/srep20167>
- Plotnikova A, Kellner MJ, Schon MA, Mosiolek M, Nodine MD (2019) MicroRNA dynamics and functions during *Arabidopsis* embryogenesis. *Plant Cell* 31:2929–2946. <https://doi.org/10.1105/tpc.19.00395>
- Qian W, Li Q, Noack J, Sass O, Meng J, Frauen M et al (2009) Heterotic patterns in rapeseed (*Brassica napus* L.): II. Crosses between European winter and Chinese semi-winter lines. *Plant Breed* 128:466–470. <https://doi.org/10.1111/j.1439-0523.2008.01597.x>
- Quezada-Martinez D, Addo Nyarko CP, Schiessl SV, Mason AS (2021) Using wild relatives and related species to build climate resilience in Brassica crops. *Theor Appl Genet* 134:1711–1728. <https://doi.org/10.1007/s00122-021-03793-3>
- Quezada-Martinez D, Zou J, Zhang W, Meng J, Batley J, Mason AS (2022) Allele segregation analysis of F1 hybrids between independent *Brassica* allohexaploid lineages. *Chromosoma*. <https://doi.org/10.1007/s00412-022-00774-3>
- Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- Ran L, Fang T, Rong H, Jiang J, Fang Y, Wang Y (2016) Analysis of cytosine methylation in early generations of resynthesized *Brassica napus*. *J Integr Agric* 15:1228–1238. [https://doi.org/10.1016/S2095-3119\(15\)61277-1](https://doi.org/10.1016/S2095-3119(15)61277-1)
- Raulusevičute I, Drabløs F, Rye MB (2020) DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation. *BMC Med Genomics* 13:6. <https://doi.org/10.1186/s12920-020-0657-6>

- Regmi R, Newman TE, Kamphuis LG, Derbyshire MC (2021) Identification of *B napus* small RNAs responsive to infection by a necrotrophic pathogen. *BMC Plant Biol* 21:366. <https://doi.org/10.1186/s12870-021-03148-6>
- Rice P, Longden I, Bleasby A (2000) EMBL: the European molecular biology open software suite. *Trends Genet* 16:276–277. [https://doi.org/10.1016/s0168-9525\(00\)02024-2](https://doi.org/10.1016/s0168-9525(00)02024-2)
- Robinson MD, Davis JM, Gordon KS (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. <https://doi.org/10.1093/bioinformatics/btp616>
- Rong H, Yang W, Zhu H, Jiang B, Jiang J, Wang Y (2021) Genomic imprinted genes in reciprocal hybrid endosperm of *Brassica napus*. *BMC Plant Biol* 21:140. <https://doi.org/10.1186/s12870-021-02908-8>
- Růžička K, Zhang M, Campilho A, Bodi Z, Kashif M, Saleh M et al (2017) Identification of factors required for m6A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *New Phytol* 215:157–172. <https://doi.org/10.1111/nph.14586>
- Sablok G, Srivastva AK, Suprasanna P, Baev V, Ralph PJ (2015) isomiRs: increasing evidences of isomiRs complexity in plant stress functional biology. *Front Plant Sci* 6:949. <https://doi.org/10.3389/fpls.2015.00949>
- Sands B, Yun S, Mendenhall AR (2021) Introns control stochastic allele expression bias. *Nat Commun* 12:6527. <https://doi.org/10.1038/s41467-021-26798-4>
- Schiessl S, Huettel B, Kuehn D, Reinhardt R, Snowdon R (2017) Post-polyploidisation morphotype diversification associates with gene copy number variation. *Sci Rep* 7:41845. <https://doi.org/10.1038/srep41845>
- Schiessl SV, Quezada-Martinez D, Orantes-Bonilla M, Snowdon RJ (2020) Transcriptomics reveal high regulatory diversity of drought tolerance strategies in a biennial oil crop. *Plant Sci* 297:110515. <https://doi.org/10.1016/j.plantsci.2020.110515>
- Scossa F, Alseekh S, Fernie AR (2021) Integrating multi-omics data for crop improvement. *J Plant Physiol* 257:153352. <https://doi.org/10.1016/j.jplph.2020.153352>
- Seifert F, Thiemann A, Grant-Downton R, Edelmann S, Rybka D, Schrag TA et al (2018a) Parental expression variation of small RNAs is negatively correlated with grain yield heterosis in a maize breeding population. *Front Plant Sci* 9:13. <https://doi.org/10.3389/fpls.2018.00013>
- Seifert F, Thiemann A, Schrag TA, Rybka D, Melchinger AE, Frisch M et al (2018b) Small RNA-based prediction of hybrid performance in maize. *BMC Genomics* 19:371. <https://doi.org/10.1186/s12864-018-4708-8>
- Serin Harmanci A, Harmanci AO, Zhou X (2020) CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data. *Nat Commun* 11:89. <https://doi.org/10.1038/s41467-019-13779-x>
- Serrato-Capuchina A, Matute DR (2018) The role of transposable elements in speciation. *Genes Basel* 9:254. <https://doi.org/10.3390/genes9050254>
- Shahzad K, Zhang X, Guo L, Qi T, Bao L, Zhang M et al (2020) Comparative transcriptome analysis between inbred and hybrids reveals molecular insights into yield heterosis of upland cotton. *BMC Plant Biol* 20:1–18. <https://doi.org/10.1186/s12870-020-02442-z>
- Shaikh AA, Chachar S, Chachar M, Ahmed N, Guan C, Zhang P (2022) Recent advances in DNA methylation and their potential breeding applications in plants. *Horticulturae* 8:562. <https://doi.org/10.3390/horticulturae8070562>
- Shen Y, Sun S, Hua S, Shen E, Ye C-Y, Cai D et al (2017) Analysis of transcriptional and epigenetic changes in hybrid vigor of allopolyploid *Brassica napus* uncovers key roles for small RNAs. *Plant J* 91:874–893. <https://doi.org/10.1111/tpj.13605>
- Shivaraj SM, Jain A, Singh A (2018) Highly preserved roles of Brassica MIR172 in polyploid Brassicas: ectopic expression of variants of Brassica MIR172 accelerates floral transition. *Mol Genet Genomics* 293:1121–1138. <https://doi.org/10.1007/s00438-018-1444-3>
- Shull GH (1914) Duplicate genes for capsule-form in *Bursa bursa-pastoris*. *Z. Ver-erbungslehre* 12:97–149. <https://doi.org/10.1007/BF01837282>
- Shull GH (1948) What is “heterosis”? *Genetics* 33:439–446. <https://doi.org/10.1093/genetics/33.5.439>
- Smit AFA and Hubley, R. (2008). *RepeatModeler Open-1.0*. <http://www.repeatmasker.org>. Accessed 2018
- Snyman MC, Solofoharivelo M-C, Souza-Richards R, Stephan D, Murray S, Burger JT (2017) The use of high-throughput small RNA sequencing reveals differentially expressed microRNAs in response to aster yellows phytoplasma-infection in *Vitis vinifera* cv. “Chardonnay.” *PLoS One* 12:e018262. <https://doi.org/10.1371/journal.pone.0182629>
- Stein A, Coriton O, Rousseau-Gueutin M, Samans B, Schiessl SV, Obermeier C et al (2017) Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in *Brassica napus*. *Plant Biotechnol J* 15:1478–1489. <https://doi.org/10.1111/pbi.12732>
- Tang X, Bian S, Tang M, Lu Q, Li S, Liu X et al (2012) MicroRNA-mediated repression of the seed maturation program during vegetative development in Arabidopsis. *PLoS Genet* 8:e1003091. <https://doi.org/10.1371/journal.pgen.1003091>
- ter Steeg EMS, Struik PC, Visser RGF, Lindhout P (2022) Crucial factors for the feasibility of commercial hybrid breeding in food crops. *Nat Plants* 8:463–473. <https://doi.org/10.1038/s41477-022-01142-w>
- Thiemann A, Meyer S, Scholten S (2009) Heterosis in plants: manifestation in early seed development and prediction approaches to assist hybrid breeding. *Chin Sci Bull* 54:2363–2375. <https://doi.org/10.1007/s11434-009-0326-9>
- Tzafirir I, Pena-Muralla R, Dickerman A, Berg M, Rogers R, Hutchens S et al (2004) Identification of genes required for embryo development in Arabidopsis. *Plant Physiol* 135:1206–1220. <https://doi.org/10.1104/pp.104.045179>
- van Hulst MHA, Paulo M-J, Kruijer W, Blankestijn-De Vries H, Kemperman B, Becker FFM et al (2018) Assessment of heterosis in two *Arabidopsis thaliana* common-reference mapping populations. *PLoS One* 13:e0205564. <https://doi.org/10.1371/journal.pone.0205564>
- Verma S, Attuluri VPS, Robert HS (2022) Transcriptional control of Arabidopsis seed development. *Planta* 255:90. <https://doi.org/10.1007/s00425-022-03870-x>
- VIB-Ugent (2021). *Venn*. <https://bioinformatics.psb.ugent.be/webtools/Venn/>. Accessed 2021
- Vitting-Seerup K, Sandelin A (2019) IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Bioinformatics* 35:4469–4471. <https://doi.org/10.1093/bioinformatics/btz247>
- Vollrath P, Chawla HS, Schiessl SV, Gabur I, Lee H, Snowdon RJ et al (2021) A novel deletion in FLOWERING LOCUS T modulates flowering time in winter oilseed rape. *Theor Appl Genet* 134:1217–1231. <https://doi.org/10.1007/s00122-021-03768-4>
- Wang M, Wang J (2022) Transcriptome and DNA methylome analyses provide insight into the heterosis in flag leaf of inter-subspecific hybrid rice. *Plant Mol Biol* 108:105–125. <https://doi.org/10.1007/s11103-021-01228-7>
- Wang L, Wang M-B, Tu J-X, Helliwell CA, Waterhouse PM, Dennis ES et al (2007) Cloning and characterization of microRNAs from

- Brassica napus*. FEBS Lett 581:3848–3856. <https://doi.org/10.1016/j.febslet.2007.07.010>
- Wang L, Wu LM, Greaves IK, Zhu A, Dennis ES, Peacock WJ (2017a) PIF4-controlled auxin pathway contributes to hybrid vigor in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A 114:E3555–E3562. <https://doi.org/10.1073/pnas.1703179114>
- Wang P, Yang C, Chen H, Song C, Zhang X, Wang D (2017b) Transcriptional basis for drought-resistance in *Brassica napus* L. Sci Rep 7:40532. <https://doi.org/10.1038/srep40532>
- Wang Z, Qiao Y, Zhang J, Shi W, Zhang J (2017c) Genome wide identification of microRNAs involved in fatty acid and lipid metabolism of *Brassica napus* by small RNA and degradome sequencing. Gene 619:61–70. <https://doi.org/10.1016/j.gene.2017.03.040>
- Wang Z, Wu X, Wu Z, An H, Yi B, Wen J et al (2018) Genome-wide DNA methylation comparison between *Brassica napus* genic male sterile line and restorer line. Int J Mol Sci 19:2689. <https://doi.org/10.3390/ijms19092689>
- Wei W, Li G, Jiang X, Wang Y, Ma Z, Niu Z et al (2018) Small RNA and degradome profiling involved in seed development and oil synthesis of *Brassica napus*. PLoS One 13:e0204998. <https://doi.org/10.1371/journal.pone.0204998>
- Wei Y, Li G, Zhang S, Zhang S, Zhang H, Sun R et al (2021) Analysis of transcriptional changes in different *Brassica napus* synthetic Allopolyploids. Genes Basel 12:82. <https://doi.org/10.3390/genes12010082>
- Weigel D, Colot V (2012) Epialleles in plant evolution. Genome Biol 13:249. <https://doi.org/10.1186/gb-2012-13-10-249>
- Wójcik AM, Nodine MD, Gaj MD (2017) miR160 and miR166/165 contribute to the LEC2-mediated auxin response involved in the somatic embryogenesis induction in *Arabidopsis*. Front Plant Sci 8:2024. <https://doi.org/10.3389/fpls.2017.02024>
- Wu J, Zhao Q, Yang Q, Liu H, Li Q, Yi X et al (2016) Comparative transcriptomic analysis uncovers the complex genetic network for resistance to *Sclerotinia sclerotiorum* in *Brassica napus*. Sci Rep 6:19007. <https://doi.org/10.1038/srep19007>
- Wu J, Lin L, Xu M, Chen P, Liu D, Sun Q et al (2018) Homoeolog expression bias and expression level dominance in resynthesized allopolyploid *Brassica napus*. BMC Genomics 19:586. <https://doi.org/10.1186/s12864-018-4966-5>
- Wu X, Liu Y, Zhang Y, Gu R (2021) Advances in research on the mechanism of heterosis in plants. Front Plant Sci 12:745726. <https://doi.org/10.3389/fpls.2021.745726>
- Xu J, Chen G, Hermanson PJ, Xu Q, Sun C, Chen W et al (2019) Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. Genome Biol 20:1–16. <https://doi.org/10.1186/s13059-019-1859-0>
- Xu K, Wu Y, Song J, Hu K, Wu Z, Wen J et al (2021) Fine mapping and identification of BnaC06.FtsH1, a lethal gene that regulates the PSII repair cycle in *Brassica napus*. Int J Mol Sci 22:2087. <https://doi.org/10.3390/ijms22042087>
- Yang M, Wang X, Ren D, Huang H, Xu M, He G et al (2017) Genomic architecture of biomass heterosis in *Arabidopsis*. Proc Natl Acad Sci 114:8101–8106. <https://doi.org/10.1073/pnas.1705423114>
- Yang K, Wen X, Mudunuri S, Varma GPS, Sablok G (2019) Diff isomiRs: large-scale detection of differential isomiRs for understanding non-coding regulated stress omics in plants. Sci Rep 9:1406. <https://doi.org/10.1038/s41598-019-38932-w>
- Yang Y, Saand MA, Huang L, Abdelaal WB, Zhang J, Wu Y et al (2021) Applications of multi-omics technologies for crop improvement. Front Plant Sci 12:563953. <https://doi.org/10.3389/fpls.2021.563953>
- Yao S, Liang F, Gill RA, Huang J, Cheng X, Liu Y et al (2020) A global survey of the transcriptome of allopolyploid *Brassica napus* based on single-molecule long-read isoform sequencing and Illumina-based RNA sequencing data. Plant J 103:843–857. <https://doi.org/10.1111/tpj.14754>
- Yi L, Liu B, Nixon PJ, Yu J, Chen F (2022) Recent advances in understanding the structural and functional evolution of FtsH proteases. Front Plant Sci 13:837528. <https://doi.org/10.3389/fpls.2022.837528>
- Yin H, Fan Z, Li X, Wang J, Liu W, Wu B et al (2016) Phylogenetic tree-informed microRNAome analysis uncovers conserved and lineage-specific miRNAs in *Camellia* during floral organ development. J Exp Bot 67:2641–2653. <https://doi.org/10.1093/jxb/erw095>
- Yin L, Zhu Z, Huang L, Luo X, Li Y, Xiao C et al (2021) DNA repair- and nucleotide metabolism-related genes exhibit differential CHG methylation patterns in natural and synthetic polyploids (*Brassica napus* L.). Hortic Res 8:142. <https://doi.org/10.1038/s41438-021-00576-1>
- Yoo M-J, Szadkowski E, Wendel JF (2013) Homoeolog expression bias and expression level dominance in allopolyploid cotton. Hered Edinb 110:171–180. <https://doi.org/10.1038/hdy.2012.94>
- Yu D, Gu X, Zhang S, Dong S, Miao H, Gebretsadik K et al (2021) Molecular basis of heterosis and related breeding strategies reveal its importance in vegetable breeding. Hortic Res 8:120. <https://doi.org/10.1038/s41438-021-00552-9>
- Zanini SF, Bayer PE, Wells R, Snowdon RJ, Batley J, Varshney RK et al (2022) Pangenomics in crop improvement—from coding structural variations to finding regulatory variants with pangenome graphs. Plant Genome 15:e20177. <https://doi.org/10.1002/tpg2.20177>
- Zemach A, Kim MY, Silva P, Rodrigues JA, Dotson B, Brooks MD et al (2010) Local DNA hypomethylation activates genes in rice endosperm. Proc Natl Acad Sci U S A 107:18729–18734. <https://doi.org/10.1073/pnas.1009695107>
- Zhang L, Peng Y, Wei X, Dai Y, Yuan D, Lu Y et al (2014a) Small RNAs as important regulators for the hybrid vigour of super-hybrid rice. J Exp Bot 65:5989–6002. <https://doi.org/10.1093/jxb/eru337>
- Zhang M, Xie S, Dong X, Zhao X, Zeng B, Chen J et al (2014b) Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize. Genome Res 24:167–176. <https://doi.org/10.1101/gr.155879.113>
- Zhang J, Wei L, Jiang J, Mason AS, Li H, Cui C et al (2018) Genome-wide identification, putative functionality and interactions between lncRNAs and miRNAs in *Brassica* species. Sci Rep 8:4960. <https://doi.org/10.1038/s41598-018-23334-1>
- Zhang Y, Liu C, Cheng H, Tian S, Liu Y, Wang S et al (2020) DNA methylation and its effects on gene expression during primary to secondary growth in poplar stems. BMC Genomics 21:1–18. <https://doi.org/10.1186/s12864-020-06902-6>
- Zhang L, He J, He H, Wu J, Li M (2021a) Genome-wide unbalanced expression bias and expression level dominance toward *Brassica oleracea* in artificially synthesized intergeneric hybrids of *Raphanobrassica*. Hortic Res 8:246. <https://doi.org/10.1038/s41438-021-00672-2>
- Zhang Q, Guan P, Zhao L, Ma M, Xie L, Li Y et al (2021b) Asymmetric epigenome maps of subgenomes reveal imbalanced transcription and distinct evolutionary trends in *Brassica napus*. Mol Plant 14:604–619. <https://doi.org/10.1016/j.molp.2020.12.020>
- Zhou P, Hirsch CN, Briggs SP, Springer NM (2019) Dynamic patterns of gene expression additivity and regulatory variation throughout maize development. Mol Plant 12:410–425. <https://doi.org/10.1016/j.molp.2018.12.015>
- Zhu A, Wang A, Zhang Y, Dennis ES, Peacock WJ, Greaves AIK (2020) Early establishment of photosynthesis and auxin biosynthesis plays a key role in early biomass heterosis in *Brassica*

- napus* (Canola) Hybrids. *Plant Cell Physiol* 61:1134–1143. <https://doi.org/10.1093/pcp/pcaa038>
- Ziegler DJ, Khan D, Kalichuk JL, Becker MG, Belmonte MF (2019) Transcriptome landscape of the early *Brassica napus* seed. *J Integr Plant Biol* 61:639–650. <https://doi.org/10.1111/jipb.12812>
- Zou J, Hu D, Mason AS, Shen X, Wang X, Wang N et al (2018) Genetic changes in a novel breeding population of *Brassica napus* synthesized from hundreds of crosses between *B. rapa* and *B. carinata*. *Plant Biotechnol J* 16:507–519. <https://doi.org/10.1111/pbi.12791>
- Zrimec J, Börlin CS, Buric F, Muhammad AS, Chen R, Siewers V et al (2020) Deep learning suggests that gene expression is encoded in all parts of a co-evolving interacting gene regulatory structure. *Nat Commun* 11:6141. <https://doi.org/10.1038/s41467-020-19921-4>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.