bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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

Mauricio Orantes-Bonilla[1†], Hao Wang[2†], HueyTyng Lee[1], Agnieszka A. Golicz[1], Dandan Hu[2], Wenwen Li[2], Jun Zou[2] & Rod J. Snowdon[1*]

[1] Department of Plant Breeding, IFZ Research Centre for Biosystems, Land Use and Nutrition, Justus Liebig University, Giessen, Germany

[2] National Key Laboratory of Crop Genetic Improvement, College of Plant Science & Technology, Huazhong Agricultural University, Wuhan, People's Republic of China

*** Correspondence:**

Corresponding author: Rod.Snowdon@agrar.uni-giessen.de

[†] These authors have contributed equally to this work

Keywords:

Gene expression, methylation, small RNA, seed development, parental dominance

Abstract:

The enhanced performance of hybrids though heterosis remains a key aspect in plant breeding; nonetheless, the transcriptomic and epigenomic mechanisms behind it are still not fully elucidated. In the present study, gene expression, small RNA abundance and genome-wide methylation patterns were evaluated in hybrids from two distant *Brassica napus* ecotypes during seed and seedling developmental stages using next generation sequencing technologies. A total of 71217, 773, 79518 and 31825 differentially expressed genes, microRNAs, small interfering RNAs and differentially methylated regions were identified respectively. Approximately 70% of the differential expression and methylation patterns observed could be explained due to parental dominance levels. Reproductive, developmental, and meiotic gene copies following transgressive and paternal dominance patterns were found through gene ontology enrichment and microRNA-target association analyses. Interestingly, maternal dominance was more prominent in hypermethylated and downregulated features during seed formation which contrasts strikingly with the general maternal gamete demethylation occurring during gametogenesis in most plant species. Linkages between methylation and gene expression allowed the identification of putative genetic epialleles with diverse pivotal biological functions. Furthermore, most differentially methylated regions, differentially expressed siRNAs and transposable elements were found near gene flanking regions that had no differential expression, hence, indicating their potential role in conserving essential genomic and transcriptomic loci across the parents and offspring.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Transcriptomic and epigenomic profiling in recent decades have developed higher yielding, disease resistant and stress tolerant crops (Scossa et al., 2021; Yang et al., 2021) that have not only expanded the genetic diversity in multiple crop species (Louwaars, 2018) but also elucidated 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 as well as to differentiate germplasm in plants. For instance, RNA sequencing (RNA-Seq) data developed though microarrays and next generation sequencing (NGS) has been used to build gene expression databases for such as Arabidopsis, rice, wheat andoilseed rape which can in turn be exploited for comparative expression studies (Petryszak et al., 2016; H. Chen et al., 2022). Small RNAs (sRNAs) derived from endogenous genomic loci or exogeneous sources are known to regulate various functions and responses in plant. Among them, microRNAs (miRNAs) and small interfering RNAs (siRNAS) have been classified and characterized into diverse databases, providing thus, a great start for further transcriptomic research (Griffiths-Jones et al., 2006; Lunardon et al., 2020). Epigenomic features, those resulting into phenotypical changes without alterations in DNA sequences, include chromatin interaction, histone modification and DNA methylation (Fitz-James and Cavalli, 2022). Genome-wide methylation analyses have found methylation level differences across plant species and determined phenotypic consequences as a result of genomic methylation (Bartels et al., 2018; Muyle et al., 2022).

Brassica napus (AACC, 2n=38), an allopolyploid oilseed crop derived from *B.rapa* (AA, 2n=20) *and B.oleracea* (CC, 2n=18), has been no exception to the advancements in the transcriptomic and epigenomic fields. Differential gene expression in *B.napus* has revealed key genes in flowering time, disease resistance and abiotic stress (Wu et al., 2016; P. Wang et al., 2017; Jian et al., 2019). Small RNAs profiling has also identified microRNA and siRNA sequences associated with pathogen response, abiotic stress and lipid metabolism in oilseed rape (Z. Wang et al., 2017; Jian et al., 2018; Martinez Palacios et al., 2019; Regmi et al., 2021). Genome-wide DNA methylation research has likewise detected methylated regions and patterns that 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).

Recent 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 in a major commercial *B. napus* hybrid that were linked to heterosis, a hybrid specific effect of high interest for crop improvement. The enhanced performance observed 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 (D. Hu et al., 2021; Quezada-Martinez et al., 2021). Nevertheless, the transcriptomic and epigenomic networks involved in heterosis have not been fully elucidated; hence limiting the potential of regulatory and non-coding features in plant breeding.

RNA-Seq and methylation-based studies have dissected putative heterotic loci in embryo and seed developmental stages in hybrid plants (Meyer et al., 2012; Kawanabe et al., 2016; Alonso-Peral et al., 2017; L. Chen et al., 2022). Early heterosis through increase of cell size and number, seed yield and biomass has been reported in *A. thaliana* and maize (Jahnke et al., 2010; L. Wang

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

et al., 2017; Zhu et al., 2020; Groszmann et al., 2014). The latter authors 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). Such parental dominances are attributed to have a main 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 leads to a genomic shock that can further alter the hybrid transcriptome (Bird et al., 2018).

Furthermore, 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 the best of our knowledge, no previous study has evaluated dominance level patterns in expression and methylation in *B.napus* hybrids directly derived from two *B.napus* parents. Therefore, the present study analyses transcriptomic and epigenomic differences during seed and seedling development in winter ecotype *B.napus* Express 617 (Lee et al., 2020) and semi-winter ecotype *B.napus* G3D001 (Zou et al., 2018) and their respective hybrid. For this purpose, mRNA, small RNA and whole-genome bisulfite sequencing were carried in all aforementioned genotypes. 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 and integration of omics features were performed to find putative interactions between found features and consequently evaluate their epigenomic and transcriptomic impact in early heterosis.

Material and Methods

Experimental design and growing conditions

Seeds from winter-type oilseed rape Express 617 (maternal line), semi-winter semi-synthetic oilseed rape G3D001 (paternal line) and their F1 offspring were planted at the same time 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 AM under liquid nitrogen. Flower buds with similar sizes were selected on the fifth day after reaching full flowering (BBCH65) to perform selfing in all genotypes and crosses between Express 617 (female recipient) and G3D001 (male donor). The newly generated F1 crosses were employed to analyze the transcriptomic and epigenomic differences during seed formation between ovules pollinated from selfed-F1 plants and those pollinated from outcrossing from Express 617 and G3D001 which from here onwards are referred as F0. Pollinated ovules were taken out with tweezers 15 and 30 days after pollination (DAP) at 10 AM and quickly placed in sampling tubes under liquid nitrogen. Biological replicates consisted of pooled-sampling 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. Moreover, three biological replicates were used for messenger and small RNA expression experiments. Two biological replicates were used for methylation studies instead, due to low material availability for some samples.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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.). All libraries were sequenced using an Illumina NovaSeq 6000 platform (Illumina Inc.). Read quality was evaluated with FastQC v.0.11.9 (Andrew S., 2010) and multiqc v.1.9 (Ewels et al., 2016) for all sequencing types.

mRNA and sRNA alignments

mRNA libraries were first filtered by selecting reads with an exact 150 bp length, 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 (Lee et al., 2020) were identified by first converting its 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 J., 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 were 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 had 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 H., 2016). Then sRNA libraries were aligned against the Express 617 reference (Lee et al., 2020) 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 based 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 siRNAs. Both miRNAs and siRNAs clusters without any coverage in all biological samples were discarded prior differential expression analysis.

Expression level dominance analysis

The gene and sRNA differential expression patterns between the hybrid and parents were assessed by comparing tissues within genotype trios in the five following stages: leaves at the six-true opened leaves stage from parents and F1 (BBCH16); 15 days after pollination ovules from selfed parents and F1 (OS15-F1) or F0 (OS15-F0); and from 30 days after pollination ovules from selfed parents and F1 (OS30-F1) or F0 (OS30-F0). Differentially expressed genes (DEGs), differentially expressed miRNAs (DE-miRNAs) and differentially expressed siRNAs

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(DE-siRNAs) between genotypes for each stage were identified using 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, dominant or transgressive. Gene expression heatmaps were generated with idep93 (Ge et al., 2018) using correlation distances and average linkages, and differentially expressed genes or sRNA 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 sugbenome, genotype and stage were calculated to evaluate subgenomic expression bias.

Gene ontology enrichment

Gene models in the Express 617 reference assembly (Lee et al., 2020) were functionally annotated through synteny comparison against the Darmor v.4.1 (Chalhoub et al., 2014) with inparanoid v.4.2 (O'Brien et al., 2005) using bootstrap, a BLOSUM80 (BLOcks SUbstitution Matrix) and an initial cut-off 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 gene ontology enrichment of biological processes based on expression level dominance for each stage, as well in comparisons between the F1 and F0 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 kept. 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 blasting their coding sequences against the Araport v.11 *Arabidopsis thaliana* coding sequences model (Cheng et al., 2017) with BLAST. Only the hit with the lowest expect value and not greater than 1.0×10^{-4}, lowest identity percentage equal or above 90% and without any opened 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

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 (Lee et al., 2020) was bisulfite converted and indexed with Bismark v.0.23 (Krueger and Andrews, 2011) *bismark_genome_preparation* tool. Filtered reads were aligned to the bisulfite converted genome using *bismark* under default settings. Duplicates were afterwards removed with *deduplicate_bismark* and methylated cytosines (mC's) were extracted using *bismark_methylation_extractor* while ignoring the first 2 basepairs from both 5' and 3' ends for both reads of a pair. Methylated C's in either the CpG, CHG or CHH methylation context were selected and converted to a browser extensible data (BED) format with *bismark2bedGrap*h 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's in every methylation context was calculated with the *coverage2cytosine* from the Bismark package. The coverage of mC's 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 DMRs methylation levels (i.e. the number of reads supporting methylation) were extracted from DMR output files and student's t-test ($p < 0.05$) were 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 following the same categorization employed for ELD by Yoo et al. (2013). Shared and unique DMR across all stages were found with the *Venn Diagrams* tools (VIB-UGent, 2021).

Cytosine methylation statistics and identification of methylated features

The number of methylated cytosines 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 carried 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 which loci coincided with corresponding additively or dominantly methylated DMRs were defined as putative genetic epialleles.

Heatmaps comparing the gene methylation and expression in transgressive DEGs were made with Heatmapper (Babicki et al., 2016) using Euclidean distances and average linkages to analyze the interaction between expression and methylation. Moreover, repeats in the Express 617 assembly were assigned to repeat families using RepeatModeler (Smit, A. F. A. and Hubley, R., 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).

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

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 with an FDR post-hoc adjustment ($p < 0.05$) were carried to find significant associations between differentially methylated and non-methylated features and distance to genes or DEGs across all stages.

Segmental expression assessment

Clustering of DEGs across chromosomal segments observed on *circlize* generated plots were further investigated. In order 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 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 kept. 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, available short read genomic 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.). Reads quality was evaluated with FastQC v.0.11.9 and libraries were afterwards 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 modified deletion-duplication pipeline previously described (Stein et al., 2017) with the exception that outliers were removed if the depth was above 100 and that deletions and duplications were defined as 25 kbp length segments that are one standard deviation above or below the mean coverage. Deletion and duplication were saved in tab-separated files and intersected with differentially expressed segments using bedtools *intersect* function.

Results

Increased maternal dominant expression and methylation during seed development

Expression and methylation patterns from Express 617, G3D001 and their F1 were compared during seed and seedling developmental stages. The parents were crossed during the experiment to evaluate developmental differences between selfed-F1plants and Express 617 x G3D001 pollinated ovules that will develop into F1 plants, which from here onwards are referred as F0 as displayed in Fig. 1. Next generation sequencing yielded abundant coverage for each

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

biological replicate as reported in Tables S1-S3. Approximately 6.8 Gbp of mRNA sequences per biological replicate were aligned against the Express 617 assembly (Lee et al., 2020) using HISAT2 v. 2.2.1 splice site aware aligner (Kim et al., 2019) consequently 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 31x genome coverage per biological replicate were aligned and processed with Bismark v.0.23 (Krueger and Andrews, 2011) as reported on Table S3.

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 71217 DEGs, 773 DE-miRNAs, 79518 DE-siRNAS and 31825 DMRs in both CpG and CHG methylation contexts were identified across all possible parents and hybrid comparisons per stage (Tables S4-S8) which were overall evenly distributed across all chromosomes (Tables S9-S13). Differential features were further classified by their expression level dominance and methylation level dominance (MLD) as specified in Materials and Methods and reported in Figure 2. More than 90% of the differentially expressed and methylated features belong to the parental dominant and additivity models. Moreover, maternal dominance accounted for approximately 89%, 85%, 83% and 60% from all detected DEGs, DE-miRNAs, DE-siRNAS and DMRs in the F0 respectively, whereas in the F1-selfed offspring the paternal dominance was more prevalent (Table S14). Furthermore, most maternal dominant DMRs in the F0 were hypermethylated, whereas DEGs were downregulated which is contrasting 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-F1 plants compared to those from the recently formed F0. Maternal dominance from Express 617 accounts for most of the DEG and DE-siRNAs patterns observed, hence indicating the maternal relevance in seed development. Interestingly no gene expression bias was found between the A and C subgenomes (Table S15, Figure S1-S5); nevertheless, more upregulation was observed in the paternal line while the maternal one displayed more downregulationduring seed development. This contrasts with the expected gene silencing in the maternal genome that can be attributed to maternal demethylation during seed formation. Moreover, a slightly higher number of differentially expressed features following maternal expression patterns were found in the F0 than in the selfed-F1; however, this might be due to the allele segregation in the selfed-F1 plants that would lead to the maternal parent being heterozygote and putatively reducing the number of features with maternal dominant expression.

The number of detected DEGs, DE-miRNAs, DE-siRNAS and DMRs in CpG and CHG contexts shared between all stages were 1565, 12, 1111, 896 and 650 accordingly (Tables S16). Altogether, differential features shared across all stages and genotypes corresponded to 3% from all found features, whereas features unique to each stage compromised approximately 2% (Table S16). Moreover, differential features that had the same dominance level patterns across sampling stages are presented in Fig. 3. Interestingly, features that had more constant dominance patterns in a higher number of stages were in the maternal dominance group. These features were mostly shared between early and late pollinated ovule stages in the F1 and F0 (Table S17-S20), indicating the relevance of the maternal genotype during seed formation for both genotypes.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DEGs and miRNAs regulate seed development

Gene ontology (GO) enrichment for biological processes was carried for all stages based on their expression level dominance. Significant enrichment for pivotal biological functions such as amino acid and carbohydrate synthesis, photosynthesis, protein transport and DNA repair and replication were found in 15 and 30 days after pollination ovules (Table S21). No significant enrichment was identified in leaves during the seedling stage. Interestingly, only transgressively upregulated genes in F1 ovules after 15 days of pollination displayed terms associated with reproduction and meiosis as shown in Figure 4, Table S22 and Figures S6-S10. Differential gene expression and gene ontology between the F1 and F0 15 days after pollination showed that the F1 engages more in photosynthesis-related functions, whereas the F0 has an increased accumulation of energy reserve compounds and cell mobilization (Table S21).

Additionally, 51 putative mRNA targets from all DE-miRNAs were detected across all stages (Table S23). Interactions between DE-miRNAs and their mRNA targets were analyzed for targets which were DEGs as reported in Table 1 and Table S24. Most DE-miRNAs associated with DEG targets had downregulated expression in the parents and F1 (ELD IX) and were more abundant during the late seed developmental stage. Expression from ortholog copies from *PHABULOSA*, *REVOLUTA* and *TOE2* (*TARGET OF EARLY ACTIVATION TAGGED 2*), which are involved in plant growth and development, was not increased despite the miRNAs targeting them being lowly expressed. Likewise, positive proportional expression interactions were observed in the *EMB2204* gene copy (Figure 5), whereas an inversely proportional relation between the miRNA and mRNA target was found for the *EMB2016* ortholog copy. Both copies are linked to embryo development yet are differently regulated, hence further research is required to elucidate their role in seed formation. *PHABULOSA*, and possibly *PHAVOLUTA*, are positive regulators of the (*LEAFY COTYLEDON 2*) gene which is a regulator of seed maturation (Tang et al., 2012). A BLAST search from the *A. thaliana* Araport 11 assembly (Cheng et al., 2017) *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 assembly (Lee et al., 2020) and it was found to be differentially expressed likewise in late seed development (Table S17 and S24) which highlights the putative broader and indirect impact from miRNAs through gene network interactions.

Methylated features in early seed formation

Methylation levels were higher in the CpG contexts with an average of 80% in all stages and genotypes as illustrated in Fig. 6, while CHH context showed the lowest methylation levels ranging from 20% to 27% despite having the highest number of methylated cytosines (Table S25, Fig S11-S14). No methylation bias per chromosome was observed (Table S25). 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 S26). 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 S27). Most differentially methylated transposable element (TE) families and superfamilies coincided with those that are more present in the Express 617 reference such as LTR (long terminal repeat) Copia and Gypsy families and approximately 70% of them were in 5 kbp gene flanking regions (Table S28-S29). Chi-square tests followed by FDR adjusted post-hoc testing at $p > 0.05$ showed that there is a significant association between the analyzed genomic features and distance to genes and DEGs (Table S30). The analyzed features included DMRs as well as differentially methylated and non-methylated DE-siRNAs and TEs. Interestingly, around 70% of features were in 5 kbp gene flanking regions; nevertheless, only 20% of them were found in 5kbp DEGs flanking regions

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Table S30). This finding, along with only 1% of all genes being differentially methylated (Table S31), points out to a putative conservative gene regulation across genotypes.

Moreover, 392 genes that were both differentially expressed and differentially methylated were regarded as putative epialleles and are listed on Table S32. Although no gene ontology enrichment was found on putative epialleles, they covered diverse biological functions such as DNA transcription, carbohydrate and lipids metabolic processes and photosynthesis (Table S33). Interestingly, both the gene body and its promoter were methylated in most putative genetic epialleles (Table S34). Most DMRs were less than 5 kbp away from a gene, hence revealing a potential regulatory role (Fig.6 and Fig. S11-S14). Both proportional and inversely proportional relationships were detected between gene methylation and gene expression in most stage comparisons (Fig. S15-S17), nevertheless, more proportional interactions in which hypomethylated genes were generally upregulated were observed during early seed development in the hybrid as shown in Figure 7.

In addition, 112635 CpG islands were detected in all assigned chromosomes in the Express 617 reference (Lee et al., 2020) with a 363 bp average length and varying concentrations in centromeric regions (Table S35). While 86% from all found CpG islands were methylated and had an average 62% methylation level, only 1.35% from them were differentially methylated (Table S36) hence displaying a putative conservatory role.

Segmental and subgenome expression bias in hybrids

Differential gene expression in genomic segments was preliminary observed through 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 were consistent DEGs patterns per segment, chromosome, genotype and stage were grouped as described in Materials and Method. Consequently, a 144 differentially expressed segments across genotypes and stages were determined (Table S37). More differentially expressed segments were found in the A subgenome and most segments found in F0 comparisons followed the same expression patterns from the maternal parent Express 617 as shown in Table S37 and in the example on chrA03 in Fig. 7. Available genomics reads from G3D001 pollinated ovules were employed to discard that the observed patterns were due to genomic rearrangements (Table S38). 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 locus (Fig. S19-S20, Table S37). However, no large-scale rearrangements were found in chromosome A03 in G3D001 (Fig. S21) as to explain the observed low expression as part of a large deletion effect in both early (15 days) and late (30 days) seed development stages (Fig. 8 and Fig. S22). Moreover, no duplication seems to be present in Express 617 as to account for its high expression, since neither the F1 nor the F0 showed a high expression pattern that could have been inherited from a large-scale duplication from the Express 617 genotype (Fig.S23-24). No specific relations were found in terms of methylation level, repeat density or relative position to centromere. This altogether leads us to hypothesize that such segment follows the maternal expression patterns more than other chromosomal loci. The mechanisms of such phenomenon could be associated to parental roles during embryo development, genomic imprinting or chromatin activity and/or genome accessibility for transcription; however, it requires further research to be elucidated.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

In the present study, features with differential expression and methylation patterns in hybrids were detected during seed development and seedling stages. Most DEGs in the hybrids showed maternal and paternal dominances regardless of the tissue and stage (Table 1). Similar parental gene expression dominance has been reported in oilseed rape and cotton (Yoo et al., 2013; Wu et al., 2018; Wei et al., 2021). Furthermore, Li et al. (2020) demonstrated that EDL can vary based on the sampled tissue, where stems and leaves showed more additive gene expression in *B. napus* when compared to the expression from its phylogenetic ancestors *B.rapa* and *B.oleracea*. Gene expression additivity was also reported by L. Zhang et al. (2021) as a main expression dominance level pattern factor in excised pod sections in crosses from *Raphanus sativus* (RR, 2n =18) and *B. oleracea*, whereas the rest of seed and pod dissected regions displayed more paternal dominances. The diversity of sampled tissues, species and genotypes in the previous studies and ours could account for the contrasting EDL observed. Differential gene expression was also observed in large segments; and was not necessarily related to genomic rearrangements (Table S37, Fig. 8). Chromatin rearrangement and transcription accessibility could be one reason explaining the large-scale differentially expressed segments observed. Unfortunately, not enough tissue material was available in all stages to discard expression bias because of genomic rearrangements since all sequencing types were carried from the same tissue. One recent alternative to overcome this issues is proposed through new structural variant callers that work directly on RNA-Seq data to call copy number variation (CNVs) without the need of genomic data (Serin Harmanci et al., 2020).

Additionally, pollinated ovules that would become F1 plants, F0 genotype, showed a higher similarity to the maternal genotype Express 617 in terms of gene expression, small RNA expression and methylation which might be due to the triploid nature of the endosperm having two maternal copies and one paternal copy through double-fertilization as discussed in Jahnke et al. (2010). Seeds are composed by a seed coat, an embryo and an endosperm, with the latter generally accounting for the largest volume in the seed. Although the separation from all seed components was out of the scope in our study; recent transcriptomic profiling of the aforementioned tissues through 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).

Subgenomic expression bias has been reported earlier in Brassica species (Bird et al., 2018; Bird et al., 2021a; Q. Zhang et al., 2021); therefore, expression bias of differentially expressed up- and downregulated genes was investigated in each genotype and stage per subgenome. Despite no subgenome bias in gene expression being detected, more genes were differentially upregulated in G3D001 than in Express 617 (Table S15, Figures S1-S5). The observed genotype-specific bias can be a result of the following genomic, transcriptomic and epigenomic factors. Firstly, genomic rearrangements such as structural variations and gene copy number variations are known to affect various traits in 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 down regulation 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 as well as difference between the parents *cis-trans* regulating factors, miRNA isoforms (isomiRs) and TE families and densities could all result into potential genotype or haplotype biased expression (Jain et al., 2018; Go and Civetta, 2020; Gill et al., 2021).

Approximately 12-18% of features shared in at least 2-3 stages followed maternal dominant patterns as shown in Fig. 3, hence, highlighting the relevance of maternal genotype in

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

transcriptomic and epigenomic scenarios. Furthermore, less than 3% from all expression and methylation features had the same expression and methylation patterns across all stages (Table S16), indicating that the role of those features might be more essential throughout seed and early seedling development. Complete lists of all features are provided in Table S17-S20 to allow further feature assessment in the context of seed development in hybrids. In addition, gene ontology enrichment showed enrichment for key biological functions involved in growth and development (Fig. 4, Table S21). Furthermore, DEGs involved in reproduction and meiotic functions were detected as having a transgressive upregulated expression in 15 days after pollination ovules in seeds from selfed-F1 plants. The list of all DEGs with GO enrichment have a potential for seed development in hybrids and are available on Supplementary Tables S21-S22.

Differentially expressed miRNAs analysis in early and late seed development showed miRNA families normally involved in plant growth and development (Plotnikova et al., 2019; Dong et al., 2022; Verma et al., 2022). miR172 regulates not only the flowering time pathway, but also embryo development through controlling of *AP2* (*APETALA 2*) and *AP2-like* genes such as *TOE2* as reported in previous studies (Boutilier et al., 2002; Shivaraj et al., 2018; Nowak et al., 2022). 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 *REVOLUTA*, *PHAVOLUTA* and *PHABULOSA* genes as discussed in Wang et al. (2007) and Tang et al. (2012). Both *PHABULOSA* and *PHAVOLUTA* 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). miR169 targets *CBF* (*C-REPEAT BINDING FACTORS*) and *NF-YA* (*NUCLEAR FACTOR YA*) genes as described in Dong et al. (2022); however, in our study, miR160 targeted an *EMB2016* (*EMBRYO DEFECTIVE 2016*) gene copy during late seed development. *EMB2016* is a member of the EMB family which are critical for embryo development (Tzafrir et al., 2004; Růžička et al., 2017; Meinke, 2020). *EMB2204* (*EMBRYO DEFECTIVE 2204*), is another EMB family member, which was targeted by miR3629 in the present study. mir3629 was first reported in *Vitis vinifera* cv. Pinot Noir by Pantaleo et al. (2010) and has been reported in *Camellia azalea*, in chilling response in *Prunus persica* as well as 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).

Differential expression from targets of the aforementioned miRNA families was observed in our studies during seed development, and further target validation 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) as well as gene co-expression networks (Schiessl et al., 2020) can prove beneficial for identifying their role in seed and embryo formation in *B.napus*. A list of differential miRNA sequences and their putative targets are found in Tables S23-S24 for further validation.

The number of methylated cytosines in the CHH context was overall higher in all genotypes compared to toher contexts; nonetheless, while methylation levels were higher in CpG and CHG contexts as also observed in multiple plants species (Niederhuth et al., 2016; Bartels et al., 2018). Methylation is generally associated with gene downregulation through transcription

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

inhibition as shown in Fig.7. Nevertheless, hypermethylation and hypomethylation were also linked with up- and downregulation respectively (Fig. S15-S17). Proportional gene hypermethylation and gene upregulation was observed in mice and human cells (Arechederra et al., 2018; Rauluseviciute et al., 2020); however, no mechanisms explaining gene activation through hypermethylation are fully known so far; hence, further research in this topic would elucidate the interactions between methylation and gene regulation.

Moreover, methylation was evaluated during seed development in our study since 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 most flowering plants like Arabidopsis, rice and maize endosperm (Gehring et al., 2009; Zemach et al., 2010; Zhang et al., 2014). Contrastingly, we observed more maternal hypermethylation and paternal hypomethylation in the F0. Similar parental methylation trends were also observed in the F1 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*. One possible explanation for this was proposed by the latter authors, where siRNAs in endosperm (siren siRNAs) were highly expressed in the seed coat and could trigger a maternal DNA methylation control in seed development through the movement of maternally expressed siRNAs from the seed coat or maternal-specific expression in the endosperm.

The molecular mechanisms and effects of genomic imprinting, where an allele follows a parental expression pattern due to inherited epigenomic modifications, is restricted mostly to the endosperm rather to the embryo in flowering plants, and has been extensively discussed in Weigel and Colot (2012), and Batista and Köhler (2020). The role of imprinted genes has been linked to chromatin modification, hormone biosynthesis, nutrient transfer, endosperm proliferation and seed size regulation as reviewed by Jiang and Köhler (2012) and Batista and Köhler (2020). Furthermore, Rong et al. (2021) reported the enrichment of transposable elements in imprinted genes in *B. napus* located in 5 kbp flanking regions. 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. It was 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 found in the seedling and seed development stages in our study (Table S32) for future confirmation through genomic sequencing and phenotypic data.

Most frequently differentially methylated transposable elements corresponded to those more abundant in the Express 617 reference (Lee et al., 2020) like the Copia and Gypsy families. Transposable elements are considered not only as key factors in speciation and subgenome expression patterns (Bird et al., 2018; Bottani et al., 2018; Bird et al., 2021b) but also as genomic features with high variability across plant species (Novák et al., 2020; Mhiri et al., 2022). In addition, comparison of transposable element densities and compositions between Express 617 and G3D001 through genome assembly of the latter would provide more insights into how TEs diverge during seed development and between cultivars.

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 as discussed in Matzke and Mosher (2014) and Gill et al. (2021). Similarly to Rong et al. (2021) most differentially methylated TEs were found in 5 kbp gene flanking regions, as expected since 78% of TEs are in such loci in the employed Express 617 reference (Lee et al., 2020). Our results also outline that most TEs, DMRs and DE-siRNAS converged in

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5 kbp gene flanking regions rather than in 5 kbp DEG flanking regions (Tables S30); thus, highlighting a putative conservation of most genetic functions by reducing the number of DMRs, DE-siRNAs and differentially methylated TEs in the proximity of DEGs. Additionally, most CpG islands were not differentially methylated between genotypes hence indicating a potential epigenomic and transcriptomic conservative role (Table S35). CpG islands are known to have diverse functions in plants (Ashikawa, 2001) and have also been analyzed in recent studies with Brassica species (Perumal et al., 2020; Park et al., 2021).

Expression and methylation dominance levels are a fast way of interpretating -omics data, and to our knowledge, the present study is the first one to analyze expression and methylation dominance level patterns in hybrids derived from two *Brassica napus* parents. Recent advances in allele specific 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 F1 plants and populations (Higgins et al., 2018; Cortijo et al., 2019; Orantes-Bonilla et al., 2022; Quezada-Martinez et al., 2022) would prove useful in improving the resolution of ELD and MLD analyses. Furthermore, multi-omics features have been employed in genomic selection in plants; hence, defining the role from each -omic feature per stage could be used to enhance expression and phenotype prediction modelling in *B.napus* (Seifert et al., 2018; Zrimec et al., 2020; Cheng et al., 2021; H. Hu et al., 2021; Knoch et al., 2021). Altogether our findings highlight the transcriptomics and epigenomic differences between early developmental stages in F1 and F0, and a similarity of the latter to Express 617 in terms of methylation level as well as in gene and small RNA expression. Differentially expressed and methylated features detected during seed development were identified and are provided in the present study for further research. Future developments in sequencing and bioinformatics will aid in elucidating the role and interactions of each transcriptomic and epigenomic feature at a higher resolution.

Data availability and statement

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.

Author contributions

RS and JZ conceived and supervised the study. MOB drafted the manuscript and designed the bioinformatic analyses. HW conducted bioinformatic studies, generated crosses, extracted and sampled pollinated ovules for sequencing and contributed to data analysis. DH carried and supervised the field experiment. WL contributed to experimental trials and HTL carried genome syntheny analyses and contributed to whole-genome bisulfite and mRNA analyses. AAG contributed in transcriptomic and epigenomic features analyses. All authors read and approved the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Funding

This research was carried under the Joint Sino-German Research (2018) frame and was sponsored by the German Research Foundation (DFG, grant number SN14/22-1 to RJS) and the National Natural Science Foundation of China (NSFC, grant number 31861133016 to JZ).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

We would like to acknowledge the support from the BMBF-funded de.NBI Cloud within the German Network for Bioinformatics Infrastructure (de.NBI) and the Bioinformatics Core Facility at JLU for the provided bioinformatic resources and support. The experimental design was created with BioRender.com.

Supplementary Information

Supplementary Figures 1–24

Supplementary Tables 1–38

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure legends

Figure 1. Transcriptomic and epigenomic experimental design. Leaves samples were taken at the six-leaves stages (BBCH16). Express 617, G3D001 and their respective F1 were selfed. Likewise, the parents were crossed during the experiment to develop pollinated ovules (F0). Pollinated ovules were sampled and sequenced 15 (OS15) and 30 (OS30) days after pollination.

Figure 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 patterrn are displayed in dot-ended lines where the dot and E, G and F abbreviations represent the Express 617, G3D001 and F1/F0 genotypes respectively. Percentages are displayed with colored backgrounds to represent high (red) or low (blue) abundance.

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

Figure 4. Gene expression heatmap (a) and gene ontology (GO) enrichment of biological processes (b) from 15 days after pollination ovules with transgressive upregulation patterns in the F1.

Figure 5. Selected normalized expression from differential expressed miRNA (DE-miRNA) and their differential expressed target gene (DEG) in 30 days after pollination ovules in the F1 and parental genotypes. a, Inversely proportional miRNA-mRNA target expression. **b**, Proportional miRNA-mRNA target expression**.**

Figure 6. Methylation patterns in 15 days after pollination ovules from F0 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.

Figure 7. Gene expression and gene methylation in CpG and CHG contexts from 15 days after pollination ovules displaying transgressive patterns in the F0 and its parents. Genes are sorted in the same order in both heatmaps.

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 8. Differentially expressed genes (DEGs) and methylation levels from 15 days after pollination ovules from F0 and parents in chromosome A03. Outer to inner tracks correspond to: **a**, predicted centromere loci in black; **b,** repeat density per 1 kbp bin; **c-e**, Express 617, F0 and G3D001 DEGs regulation; **f-h,** Express 617, F0 and G3D001 methylation levels per 1 kbp bin. Differential expression segment starting at approximately 11 Mbp to 18.8 Mbp is highlighted in orange

Table 1. Predicted mRNA target from differentially expressed miRNAs (DE-miRNAs) in 15 and 30 days after pollination ovules in F1 and parents by expression level dominance (ELD).

Stage	Predicted DEmiRNA familiy	DEmiRNA-ELD	miRNA-Target	miRNA-target-ELD	*A. thaliana* homolog ID	*A. thaliana* homolog name
OS15-F1	miRNA 165/166 A miRNA 165/166 B miRNA 165/166 C	IX	C04p041520.1_BnaEXP	IX	AT2G34710	PHABULOSA
OS30-F1	miRNA 3629 A	VIII	A02p005120.1_BnaEXP	VIII	AT1G22090	EMB2204
	miRNA 166 A miRNA 165/166 B miRNA 166 B miRNA 166 C	IX	C04p041520.1_BnaEXP	IX	AT2G34710	PHABULOSA
	miRNA 166 A miRNA 165/166 B miRNA 166 B miRNA 166 C	IX	C05p024820.1_BnaEXP	XII	AT1G30490	PHAVOLUTA
	miRNA 9410/9411 A miRNA 9410 A	IX	A06p017700.1_BnaEXP	VII	AT3G47060	FTSH PROTEASE 7
	miRNA 165/166 B	IX	A02p009610.1_BnaEXP	IX	AT5G60690	REVOLUTA
	miRNA 169 A	IX	A03p029900.3_BnaEXP	XI	AT3G05680	EMB2016
	miRNA 172 A	IX	C09p033660.1_BnaEXP	IX	AT5G60120	TOE2

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Alonso-Peral, M. M., Trigueros, M., Sherman, B., Ying, H., Taylor, J. M., Peacock, W. J., et al. (2017). Patterns of gene expression in developing embryos of Arabidopsis hybrids. *Plant J* 89, 927–939. doi: 10.1111/tpj.13432

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Andrew 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, S. K., 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. doi: 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. doi: 10.1046/j.1365-313x.2001.01062.x

Axtell, M. J. (2013). Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol* 64, 137–159. doi: 10.1146/annurev-arplant-050312-120043

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A., et al. (2016). Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res 44, W147-53. doi: 10.1093/nar/gkw419

Barakat, A., Sriram, A., Park, J., Zhebentyayeva, T., Main, D., and Abbott, A. (2012). Genome wide identification of chilling responsive microRNAs in Prunus persica. *BMC Genomics* 13, 481. doi: 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. doi: 10.3390/ijms19072144

Batista, R. A., and Köhler, C. (2020). Genomic imprinting in plants-revisiting existing models. *Genes Dev* 34, 24–36. doi: 10.1101/gad.332924.119

Batista, R. A., Moreno-Romero, J., Qiu, Y., van Boven, J., Santos-González, J., Figueiredo, D. D., et al. (2019). The MADS-box transcription factor PHERES1 controls imprinting in the endosperm by binding to domesticated transposons. *Elife* 8. doi: 10.7554/eLife.50541

Bird, K. A., Niederhuth, C. E., Ou, S., Gehan, M., Pires, J. C., Xiong, Z., et al. (2021a). Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid Brassica napus. *New Phytol* 230, 354–371. doi: 10.1111/nph.17137

Bird, K. A., Pires, J. C., VanBuren, R., Xiong, Z., and Edger, P. P. (2021b). *Gene dosage constraints affect the transcriptional response to allopolyploidy and homoeologous exchange in resynthesized Brassica napus*.

Bird, K. A., VanBuren, R., Puzey, J. R., and Edger, P. P. (2018). The causes and consequences of subgenome dominance in hybrids and recent polyploids. *New Phytol* 220, 87–93. doi: 10.1111/nph.15256

Bottani, S., Zabet, N. R., Wendel, J. F., and Veitia, R. A. (2018). Gene Expression Dominance in Allopolyploids: Hypotheses and Models. *Trends Plant Sci* 23, 393–402. doi: 10.1016/j.tplants.2018.01.002

Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., et al. (2018). Ten things you should know about transposable elements. *Genome Biol* 19, 199. doi: 10.1186/s13059-018-1577-z

Boutilier, K., Offringa, R., Sharma, V. K., 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. doi: 10.1105/tpc.001941

Braybrook, S. A., Stone, S. L., Park, S., Bui, A. Q., Le, B. H., Fischer, R. L., et al. (2006). Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

embryo maturation and somatic embryogenesis. *Proc Natl Acad Sci U S A* 103, 3468–3473. doi: 10.1073/pnas.0511331103

Cao, L., Yu, N., Li, J., Qi, Z., Wang, D., and Chen, L. (2016). Heritability and Reversibility of DNA Methylation Induced by in vitro Grafting between Brassica juncea and B. oleracea. *Sci Rep* 6, 27233. doi: 10.1038/srep27233

Cao, S., Wang, L., Han, T., Ye, W., Liu, Y., Sun, Y., et al. (2022). Small RNAs mediate transgenerational inheritance of genome-wide trans-acting epialleles in maize. *Genome Biol* 23, 53. doi: 10.1186/s13059-022-02614-0

Catoni, M., Tsang, J. M., Greco, A. P., and Zabet, N. R. (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. doi: 10.1093/nar/gky602

Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., et al. (2014). Plant genetics. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. *Science* 345, 950–953. doi: 10.1126/science.1253435

H. Chen, T. Wang, X. He, X. Cai, R. Lin, J. Liang, et al. (2022). BRAD V3.0: an upgraded Brassicaceae database. *Nucleic Acids Res* 50, D1432-D1441. doi: 10.1093/nar/gkab1057

L. Chen, Y. Zhu, X. Ren, D. Yao, Y. Song, S. Fan, et al. (2022). Heterosis and Differential DNA Methylation in Soybean Hybrids and Their Parental Lines. *Plants (Basel)* 11. doi: 10.3390/plants11091136

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884-i890. doi: 10.1093/bioinformatics/bty560

Cheng, C.-Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., and Town, C. D. (2017). Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *Plant J* 89, 789–804. doi: 10.1111/tpj.13415

Cheng, C.-Y., Li, Y., Varala, K., Bubert, J., Huang, J., Kim, G. J., et al. (2021). Evolutionarily informed machine learning enhances the power of predictive gene-to-phenotype relationships. *Nat Commun* 12, 5627. doi: 10.1038/s41467-021-25893-w

Cortijo, S., Aydin, Z., Ahnert, S., and Locke, J. C. (2019). Widespread inter-individual gene expression variability in Arabidopsis thaliana. *Mol Syst Biol* 15, e8591. doi: 10.15252/msb.20188591

Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., et al. (2022). Ensembl 2022. *Nucleic Acids Res* 50, D988-D995. doi: 10.1093/nar/gkab1049

Dai, X., Zhuang, Z., and Zhao, P. X. (2018). psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res* 46, W49-W54. doi: 10.1093/nar/gky316

Dainat J. (2019). *AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format. (Version v0.5.0).* https://www.doi.org/10.5281/zenodo.3552717 (Accessed 2020)

Dong, Q., Hu, B., and Zhang, C. (2022). microRNAs and Their Roles in Plant Development. *Front Plant Sci* 13, 824240. doi: 10.3389/fpls.2022.824240

Doyle, J. J., and Doyle, J. (1987). A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue. *Phytochemical Bulletin* 19, 11–15.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. doi: 10.1093/bioinformatics/btw354

Fan, J., Hu, J., Xue, C., Zhang, H., Susztak, K., Reilly, M. P., et al. (2020). ASEP: Gene-based detection of allele-specific expression across individuals in a population by RNA sequencing. *PLoS Genet* 16, e1008786. doi: 10.1371/journal.pgen.1008786

Feschotte, C. (2004). Merlin, a new superfamily of DNA transposons identified in diverse animal genomes and related to bacterial IS1016 insertion sequences. *Mol Biol Evol* 21, 1769–1780. doi: 10.1093/molbev/msh188

Fitz-James, M. H., and Cavalli, G. (2022). Molecular mechanisms of transgenerational epigenetic inheritance. *Nat Rev Genet* 23, 325–341. doi: 10.1038/s41576-021-00438-5

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fujimoto, R., Uezono, K., Ishikura, S., Osabe, K., Peacock, W. J., and Dennis, E. S. (2018). Recent research on the mechanism of heterosis is important for crop and vegetable breeding systems. *Breed Sci* 68, 145–158. doi: 10.1270/jsbbs.17155

Gao, C., Zhou, G., Ma, C., Zhai, W., Zhang, T., Liu, Z., et al. (2016). Helitron-like transposons contributed to the mating system transition from out-crossing to self-fertilizing in polyploid Brassica napus L. *Sci Rep* 6, 33785. doi: 10.1038/srep33785

Gardiner-Garden, M., and Frommer, M. (1987). CpG Islands in vertebrate genomes. *Journal of Molecular Biology* 196, 261–282. doi: 10.1016/0022-2836(87)90689-9

Ge, S. X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* 36, 2628–2629. doi: 10.1093/bioinformatics/btz931

Ge, S. X., Son, E. W., and Yao, R. (2018). iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* 19, 534. doi: 10.1186/s12859-018-2486-6

Gehring, M., Bubb, K. L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science* 324, 1447–1451. doi: 10.1126/science.1171609

German, M. A., 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. doi: 10.1038/nbt1417

Gill, R. A., Scossa, F., King, G. J., Golicz, A. A., Tong, C., Snowdon, R. J., et al. (2021). On the Role of Transposable Elements in the Regulation of Gene Expression and Subgenomic Interactions in Crop Genomes. *Critical Reviews in Plant Sciences* 40, 157–189. doi: 10.1080/07352689.2021.1920731

Go, A. C., and Civetta, A. (2020). Hybrid Incompatibilities and Transgressive Gene Expression Between Two Closely Related Subspecies of Drosophila. *Front Genet* 11, 599292. doi: 10.3389/fgene.2020.599292

Golicz, A. A., Allu, A. D., Li, W., Lohani, N., Singh, M. B., and Bhalla, P. L. (2021). A dynamic intron retention program regulates the expression of several hundred genes during pollen meiosis. *Plant Reprod* 34, 225–242. doi: 10.1007/s00497-021-00411-6

Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., and Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 34, D140-4. doi: 10.1093/nar/gkj112

Groszmann, M., Gonzalez-Bayon, R., Greaves, I. K., Wang, L., Huen, A. K., Peacock, W. J., et al. (2014). Intraspecific Arabidopsis hybrids show different patterns of heterosis despite the close relatedness of the parental genomes. *Plant Physiol* 166, 265–280. doi: 10.1104/pp.114.243998

Grover, J. W., Burgess, D., Kendall, T., Baten, A., Pokhrel, S., King, G. J., 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. doi: 10.1073/pnas.2001332117

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize Implements and enhances circular visualization in R. *Bioinformatics* 30, 2811–2812. doi: 10.1093/bioinformatics/btu393

Higgins, E. E., Clarke, W. E., Howell, E. C., Armstrong, S. J., and Parkin, I. A. P. (2018). Detecting de Novo Homoeologous Recombination Events in Cultivated Brassica napus Using a Genome-Wide SNP Array. *G3 (Bethesda)* 8, 2673–2683. doi: 10.1534/g3.118.200118

D. Hu, J. Jing, R. J. Snowdon, A. S. Mason, J. Shen, J. Meng, et al. (2021). Exploring the gene pool of Brassica napus by genomics-based approaches. *Plant Biotechnol J* 19, 1693–1712. doi: 10.1111/pbi.13636

H. Hu, M. T. Campbell, T. H. Yeats, X. Zheng, D. E. Runcie, G. Covarrubias-Pazaran, et al. (2021). Multi-omics prediction of oat agronomic and seed nutritional traits across

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

environments and in distantly related populations. *Theor Appl Genet* 134, 4043–4054. doi: 10.1007/s00122-021-03946-4

Jahnke, S., Sarholz, B., Thiemann, A., Kühr, V., Gutiérrez-Marcos, J. F., Geiger, H. H., 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. doi: 10.1007/s00122-009-1207-y

Jain, A., Anand, S., Singh, N. K., and 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. doi: 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. doi: 10.3390/ijms19051431

Jian, H., Zhang, A., Ma, J., Wang, T., Yang, B., Shuang, L. S., et al. (2019). Joint QTL mapping and transcriptome sequencing analysis reveal candidate flowering time genes in Brassica napus L. *BMC Genomics* 20, 21. doi: 10.1186/s12864-018-5356-8

Jiang, H., and Köhler, C. (2012). Evolution, function, and regulation of genomic imprinting in plant seed development. *J Exp Bot* 63, 4713–4722. doi: 10.1093/jxb/ers145

Jiang, N., Bao, Z., Zhang, X., Eddy, S. R., and Wessler, S. R. (2004). Pack-MULE transposable elements mediate gene evolution in plants. *Nature* 431, 569–573. doi: 10.1038/nature02953

Johnson, N. R., Yeoh, J. M., Coruh, C., and Axtell, M. J. (2016). Improved Placement of Multi-mapping Small RNAs. *G3 (Bethesda)* 6, 2103–2111. doi: 10.1534/g3.116.030452

Kawanabe, T., Ishikura, S., Miyaji, N., Sasaki, T., Wu, L. M., 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. doi: 10.1073/pnas.1613372113

Khan, D., Ziegler, D. J., Kalichuk, J. L., 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. doi: 10.1111/tpj.15587

Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 37, 907–915. doi: 10.1038/s41587-019-0201-4

Kirkbride, R. C., Lu, J., Zhang, C., Mosher, R. A., Baulcombe, D. C., and Chen, Z. J. (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. doi: 10.1073/pnas.1807621116

Knoch, D., Werner, C. R., Meyer, R. C., Riewe, D., Abbadi, A., Lücke, S., et al. (2021). Multi-omics-based prediction of hybrid performance in canola. *Theor Appl Genet* 134, 1147–1165. doi: 10.1007/s00122-020-03759-x

Krueger, F., and Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 27, 1571–1572. doi: 10.1093/bioinformatics/btr167

Krueger, F., James, F., Ewels, P., Afyounian, E., and Schuster-Boeckler B. (2021). *FelixKrueger/TrimGalore: v0.6.7.* Zenodo. Available at: https://doi.org/10.5281/zenodo.5127899

Lauss, K., Wardenaar, R., Oka, R., van Hulten, M. H. A., Guryev, V., Keurentjes, J. J. B., et al. (2018). Parental DNA Methylation States Are Associated with Heterosis in Epigenetic Hybrids. *Plant Physiol* 176, 1627–1645. doi: 10.1104/pp.17.01054

Lee, H., Chawla, H. S., Obermeier, C., Dreyer, F., Abbadi, A., and Snowdon, R. (2020). Chromosome-Scale Assembly of Winter Oilseed Rape Brassica napus. *Front Plant Sci* 11, 496. doi: 10.3389/fpls.2020.00496

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 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. doi: 10.1093/bioinformatics/btp352

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. *Sci Rep* 6, 38401. doi: 10.1038/srep38401

Li, M., Wang, R., Wu, X., and Wang, J. (2020). Homoeolog expression bias and expression level dominance (ELD) in four tissues of natural allotetraploid Brassica napus. *BMC Genomics* 21, 330. doi: 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. doi: 10.3389/fpls.2021.734419

Li H. (2016). *Seqtk*. https://github.com/lh3/seqtk (Accessed 2019)

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930. doi: 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. doi: 10.1093/dnares/dsy030

Liu, J., Zhou, R., Wang, W., Wang, H., Qiu, Y., Raman, R., et al. (2020). A copia-like retrotransposon insertion in the upstream region of the SHATTERPROOF1 gene, BnSHP1.A9, is associated with quantitative variation in pod shattering resistance in oilseed rape. *J Exp Bot* 71, 5402–5413. doi: 10.1093/jxb/eraa281

Lopes, A. L. K., Kriegová, E., Lukeš, J., Krieger, M. A., and Ludwig, A. (2021). Distribution of Merlin in eukaryotes and first report of DNA transposons in kinetoplastid protists. *PLoS One* 16, e0251133. doi: 10.1371/journal.pone.0251133

Louwaars, N. P. (2018). Plant breeding and diversity: A troubled relationship? *Euphytica* 214, 114. doi: 10.1007/s10681-018-2192-5

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550. doi: 10.1186/s13059-014-0550-8

Lukasik, A., Pietrykowska, H., Paczek, L., Szweykowska-Kulinska, Z., and Zielenkiewicz, P. (2013). High-throughput sequencing identification of novel and conserved miRNAs in the Brassica oleracea leaves. *BMC Genomics* 14, 801. doi: 10.1186/1471-2164-14-801

Lunardon, A., Johnson, N. R., 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. doi: 10.1101/gr.256750.119

Mahmoud, M., Gobet, N., Cruz-Dávalos, D. I., Mounier, N., Dessimoz, C., and Sedlazeck, F. J. (2019). Structural variant calling: the long and the short of it. *Genome Biol* 20, 246. doi: 10.1186/s13059-019-1828-7

Makhoul, M., Chawla, H. S., Wittkop, B., Stahl, A., Voss-Fels, K. P., Zetzsche, H., et al. (2022). Long-Amplicon Single-Molecule Sequencing Reveals Novel, Trait-Associated Variants of VERNALIZATION1 Homoeologs in Hexaploid Wheat. *Front Plant Sci* 13. doi: 10.3389/fpls.2022.942461

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. doi: 10.1093/molbev/msz007

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* 15, 394–408. doi: 10.1038/nrg3683

Meinke, D. W. (2020). Genome-wide identification of EMBRYO-DEFECTIVE (EMB) genes required for growth and development in Arabidopsis. *New Phytol* 226, 306–325. doi: 10.1111/nph.16071

Meyer, R. C., 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. doi: 10.1111/j.1365-313X.2012.05021.x

Mhiri, C., Borges, F., and Grandbastien, M.-A. (2022). Specificities and Dynamics of Transposable Elements in Land Plants. *Biology (Basel)* 11. doi: 10.3390/biology11040488

Miller, C., Wells, R., McKenzie, N., Trick, M., Ball, J., Fatihi, 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. doi: 10.1105/tpc.18.00577

Montgomery, S. A., and Berger, F. (2021). The evolution of imprinting in plants: beyond the seed. *Plant Reprod* 34, 373–383. doi: 10.1007/s00497-021-00410-7

Morin, R. D., Aksay, G., Dolgosheina, E., Ebhardt, H. A., Magrini, V., Mardis, E. R., et al. (2008). Comparative analysis of the small RNA transcriptomes of Pinus contorta and Oryza sativa. *Genome Res* 18, 571–584. doi: 10.1101/gr.6897308

Muyle, A. M., Seymour, D. K., Lv, Y., Huettel, B., and Gaut, B. S. (2022). Gene Body Methylation in Plants: Mechanisms, Functions, and Important Implications for Understanding Evolutionary Processes. *Genome Biol Evol* 14. doi: 10.1093/gbe/evac038

Niederhuth, C. E., Bewick, A. J., Ji, L., Alabady, M. S., Kim, K. D., Li, Q., et al. (2016). Widespread natural variation of DNA methylation within angiosperms. *Genome Biol* 17, 194. doi: 10.1186/s13059-016-1059-0

Novák, P., Guignard, M. S., Neumann, P., Kelly, L. J., Mlinarec, J., Koblížková, A., et al. (2020). Repeat-sequence turnover shifts fundamentally in species with large genomes. *Nat Plants* 6, 1325–1329. doi: 10.1038/s41477-020-00785-x

Nowak, K., Morończyk, J., Grzyb, M., Szczygieł-Sommer, A., and Gaj, M. D. (2022). miR172 Regulates WUS during Somatic Embryogenesis in Arabidopsis via AP2. *Cells* 11. doi: 10.3390/cells11040718

O'Brien, K. P., Remm, M., and Sonnhammer, E. L. L. (2005). Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res* 33, D476-80. doi: 10.1093/nar/gki107

Orantes-Bonilla, M., Makhoul, M., Lee, H., Chawla, H. S., Vollrath, P., Langstroff, A., et al. (2022). Frequent spontaneous structural rearrangements promote transgenerational genome diversification in Brassica napus. Biorxiv. Available at: https://doi.org/10.1101/2022.06.27.497715

Pantaleo, V., Szittya, 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. doi: 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. doi: 10.1038/srep20167

Park, J.-S., Shin, Y.-H., and Park, Y.-D. (2021). DNA Methylation Level Changes in Transgenic Chinese Cabbage (Brassica rapa ssp. pekinensis) Plants and Their Effects on Corresponding Gene Expression Patterns. *Genes (Basel)* 12. doi: 10.3390/genes12101563

Perumal, S., Koh, C. S., Jin, L., Buchwaldt, M., Higgins, E. E., Zheng, C., et al. (2020). A high-contiguity Brassica nigra genome localizes active centromeres and defines the ancestral Brassica genome. *Nat Plants* 6, 929–941. doi: 10.1038/s41477-020-0735-y

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Petryszak, R., Keays, M., Tang, Y. A., Fonseca, N. A., Barrera, E., Burdett, T., et al. (2016). Expression Atlas update--an integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res* 44, D746-52. doi: 10.1093/nar/gkv1045

Plotnikova, A., Kellner, M. J., Schon, M. A., Mosiolek, M., and Nodine, M. D. (2019). MicroRNA Dynamics and Functions During Arabidopsis Embryogenesis. *Plant Cell* 31, 2929–2946. doi: 10.1105/tpc.19.00395

Quesneville, H. (2020). Twenty years of transposable element analysis in the Arabidopsis thaliana genome. *Mob DNA* 11, 28. doi: 10.1186/s13100-020-00223-x

Quezada-Martinez, D., Addo Nyarko, C. P., Schiessl, S. V., and Mason, A. S. (2021). Using wild relatives and related species to build climate resilience in Brassica crops. *Theor Appl Genet* 134, 1711–1728. doi: 10.1007/s00122-021-03793-3

Quezada-Martinez, D., Zou, J., Zhang, W., Meng, J., Batley, J., and Mason, A. S. (2022). Allele segregation analysis of F1 hybrids between independent Brassica allohexaploid lineages. *Chromosoma*. doi: 10.1007/s00412-022-00774-3

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033

Raizada, M. N., Benito, M.-I., and Walbot, V. (2001). The MuDR transposon terminal inverted repeat contains a complex plant promoter directing distinct somatic and germinal programs. *Plant J* 25, 79–91. doi: 10.1111/j.1365-313X.2001.00939.x

Ramakrishnan, M., Satish, L., Kalendar, R., Narayanan, M., Kandasamy, S., Sharma, A., et al. (2021). The Dynamism of Transposon Methylation for Plant Development and Stress Adaptation. *Int J Mol Sci* 22. doi: 10.3390/ijms222111387

Ran, L., Fang, T., Rong, H., Jiang, J., Fang, Y., and Wang, Y. (2016). Analysis of cytosine methylation in early generations of resynthesized Brassica napus. *Journal of Integrative Agriculture* 15, 1228–1238. doi: 10.1016/S2095-3119(15)61277-1

Rauluseviciute, I., Drabløs, F., and Rye, M. B. (2020). DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation. *BMC Med Genomics* 13, 6. doi: 10.1186/s12920-020-0657-6

Regmi, R., Newman, T. E., Kamphuis, L. G., and Derbyshire, M. C. (2021). Identification of B. napus small RNAs responsive to infection by a necrotrophic pathogen. *BMC Plant Biol* 21, 366. doi: 10.1186/s12870-021-03148-6

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics* 16, 276–277. doi: 10.1016/s0168-9525(00)02024-2

Rong, H., Yang, W., Zhu, H., Jiang, B., Jiang, J., and Wang, Y. (2021). Genomic imprinted genes in reciprocal hybrid endosperm of Brassica napus. *BMC Plant Biol* 21, 140. doi: 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 m6 A mRNA methylation in Arabidopsis reveals a role for the conserved E3 ubiquitin ligase HAKAI. *New Phytol* 215, 157–172. doi: 10.1111/nph.14586

Sablok, G., Srivastva, A. K., Suprasanna, P., Baev, V., and Ralph, P. J. (2015). isomiRs: Increasing Evidences of isomiRs Complexity in Plant Stress Functional Biology. *Front Plant Sci* 6, 949. doi: 10.3389/fpls.2015.00949

Sands, B., Yun, S., and Mendenhall, A. R. (2021). Introns control stochastic allele expression bias. *Nat Commun* 12, 6527. doi: 10.1038/s41467-021-26798-4.

Schiessl, S., Huettel, B., Kuehn, D., Reinhardt, R., and Snowdon, R. (2017). Post-polyploidisation morphotype diversification associates with gene copy number variation. *Sci Rep* 7, 41845. doi: 10.1038/srep41845

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Schiessl, S. V., Quezada-Martinez, D., Orantes-Bonilla, M., and Snowdon, R. J. (2020). Transcriptomics reveal high regulatory diversity of drought tolerance strategies in a biennial oil crop. *Plant Sci* 297, 110515. doi: 10.1016/j.plantsci.2020.110515

Scossa, F., Alseekh, S., and Fernie, A. R. (2021). Integrating multi-omics data for crop improvement. *J Plant Physiol* 257, 153352. doi: 10.1016/j.jplph.2020.153352

Seifert, F., Thiemann, A., Schrag, T. A., Rybka, D., Melchinger, A. E., Frisch, M., et al. (2018). Small RNA-based prediction of hybrid performance in maize. *BMC Genomics* 19, 371. doi: 10.1186/s12864-018-4708-8

Serin Harmanci, A., Harmanci, A. O., and Zhou, X. (2020). CaSpER identifies and visualizes CNV events by integrative analysis of single-cell or bulk RNA-sequencing data. *Nat Commun* 11, 89. doi: 10.1038/s41467-019-13779-x

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. doi: 10.1111/tpj.13605

Shivaraj, S. M., Jain, A., and 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. doi: 10.1007/s00438-018-1444-3

Sigman, M. J., and Slotkin, R. K. (2016). The First Rule of Plant Transposable Element Silencing: Location, Location, Location. *Plant Cell* 28, 304–313. doi: 10.1105/tpc.15.00869

Smit, A. F. A. and Hubley, R. (2008). *RepeatModeler Open-1.0*. http://www.repeatmasker.org (Accessed 2018)

Snyman, M. C., Solofoharivelo, M.-C., Souza-Richards, R., Stephan, D., Murray, S., and Burger, J. T. (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, e0182629. doi: 10.1371/journal.pone.0182629

Stein, A., Coriton, O., Rousseau-Gueutin, M., Samans, B., Schiessl, S. V., Obermeier, C., et al. (2017). Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in Brassica napus. *Plant Biotechnol J* 15, 1478–1489. doi: 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. doi: 10.1371/journal.pgen.1003091

Tzafrir, 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. doi: 10.1104/pp.104.045179

Verma, S., Attuluri, V. P. S., and Robert, H. S. (2022). Transcriptional control of Arabidopsis seed development. *Planta* 255, 90. doi: 10.1007/s00425-022-03870-x

VIB-UGent (2021). *Venn*. https://bioinformatics.psb.ugent.be/webtools/Venn (Accessed 2021)

Vitting-Seerup, K., and Sandelin, A. (2019). IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Bioinformatics* 35, 4469–4471. doi: 10.1093/bioinformatics/btz247

Vollrath, P., Chawla, H. S., Schiessl, S. V., Gabur, I., Lee, H., Snowdon, R. J., et al. (2021). A novel deletion in FLOWERING LOCUS T modulates flowering time in winter oilseed rape. *Theor Appl Genet* 134, 1217–1231. doi: 10.1007/s00122-021-03768-4

Wang, L., Wang, M.-B., Tu, J.-X., Helliwell, C. A., Waterhouse, P. M., Dennis, E. S., et al. (2007). Cloning and characterization of microRNAs from Brassica napus. *FEBS Lett* 581, 3848–3856. doi: 10.1016/j.febslet.2007.07.010

L. Wang, L. M. Wu, I. K. Greaves, A. Zhu, E. S. Dennis, and W. J. Peacock (2017). PIF4-controlled auxin pathway contributes to hybrid vigor in Arabidopsis thaliana. *Proc Natl Acad Sci U S A* 114, E3555-E3562. doi: 10.1073/pnas.1703179114

P. Wang, C. Yang, H. Chen, C. Song, X. Zhang, and D. Wang (2017). Transcriptomic basis for drought-resistance in Brassica napus L. *Sci Rep* 7, 40532. doi: 10.1038/srep40532

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Z. Wang, Y. Qiao, J. Zhang, W. Shi, and J. Zhang (2017). 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. doi: 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. doi: 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. doi: 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. doi: 10.3390/genes12010082

Weigel, D., and Colot, V. (2012). Epialleles in plant evolution. *Genome Biol* 13, 249. doi: 10.1186/gb-2012-13-10-249

Wójcik, A. M., Nodine, M. D., and Gaj, M. D. (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. doi: 10.3389/fpls.2017.02024

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. doi: 10.1186/s12864-018-4966-5

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. doi: 10.1038/srep19007

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. doi: 10.3390/ijms22042087

Yaakov, B., and Kashkush, K. (2011). Massive alterations of the methylation patterns around DNA transposons in the first four generations of a newly formed wheat allohexaploid. *Genome* 54, 42–49. doi: 10.1139/G10-091

Yang, K., Wen, X., Mudunuri, S., Varma, G. P. S., and Sablok, G. (2019). Diff isomiRs: Large-scale detection of differential isomiRs for understanding non-coding regulated stress omics in plants. *Sci Rep* 9, 1406. doi: 10.1038/s41598-019-38932-w

Yang, Y., Saand, M. A., Huang, L., Abdelaal, W. B., Zhang, J., Wu, Y., et al. (2021). Applications of Multi-Omics Technologies for Crop Improvement. *Front Plant Sci* 12, 563953. doi: 10.3389/fpls.2021.563953

Yao, S., Liang, F., Gill, R. A., 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. doi: 10.1111/tpj.14754

Yi, L., Liu, B., Nixon, P. J., Yu, J., and Chen, F. (2022). Recent Advances in Understanding the Structural and Functional Evolution of FtsH Proteases. *Front Plant Sci* 13, 837528. doi: 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. doi: 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. doi: 10.1038/s41438-021-00576-1

bioRxiv preprint doi: https://doi.org/10.1101/2022.09.05.506610; this version posted September 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Yoo, M.-J., Szadkowski, E., and Wendel, J. F. (2013). Homoeolog expression bias and expression level dominance in allopolyploid cotton. *Heredity (Edinb)* 110, 171–180. doi: 10.1038/hdy.2012.94

Zanini, S. F., Bayer, P. E., Wells, R., Snowdon, R. J., Batley, J., Varshney, R. K., et al. (2022). Pangenomics in crop improvement-from coding structural variations to finding regulatory variants with pangenome graphs. *Plant Genome* 15, e20177. doi: 10.1002/tpg2.20177

Zemach, A., Kim, M. Y., Silva, P., Rodrigues, J. A., Dotson, B., Brooks, M. D., et al. (2010). Local DNA hypomethylation activates genes in rice endosperm. *Proc Natl Acad Sci U S A* 107, 18729–18734. doi: 10.1073/pnas.1009695107

Zhang, J., Liu, Y., Xia, E.-H., Yao, Q.-Y., Liu, X.-D., and Gao, L.-Z. (2015). Autotetraploid rice methylome analysis reveals methylation variation of transposable elements and their effects on gene expression. *Proc Natl Acad Sci U S A* 112, E7022-9. doi: 10.1073/pnas.1515170112

Zhang, J., Wei, L., Jiang, J., Mason, A. S., 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. doi: 10.1038/s41598-018-23334-1

L. Zhang, J. He, H. He, J. Wu, and M. Li (2021). Genome-wide unbalanced expression bias and expression level dominance toward Brassica oleracea in artificially synthesized intergeneric hybrids of Raphanobrassica. *Hortic Res* 8, 246. doi: 10.1038/s41438-021-00672-2

Zhang, M., Xie, S., Dong, X., Zhao, X., Zeng, B., Chen, J., et al. (2014). Genome-wide high resolution parental-specific DNA and histone methylation maps uncover patterns of imprinting regulation in maize. *Genome Res* 24, 167–176. doi: 10.1101/gr.155879.113

Q. Zhang, P. Guan, L. Zhao, M. Ma, L. Xie, Y. Li, et al. (2021). Asymmetric epigenome maps of subgenomes reveal imbalanced transcription and distinct evolutionary trends in Brassica napus. *Mol Plant* 14, 604–619. doi: 10.1016/j.molp.2020.12.020

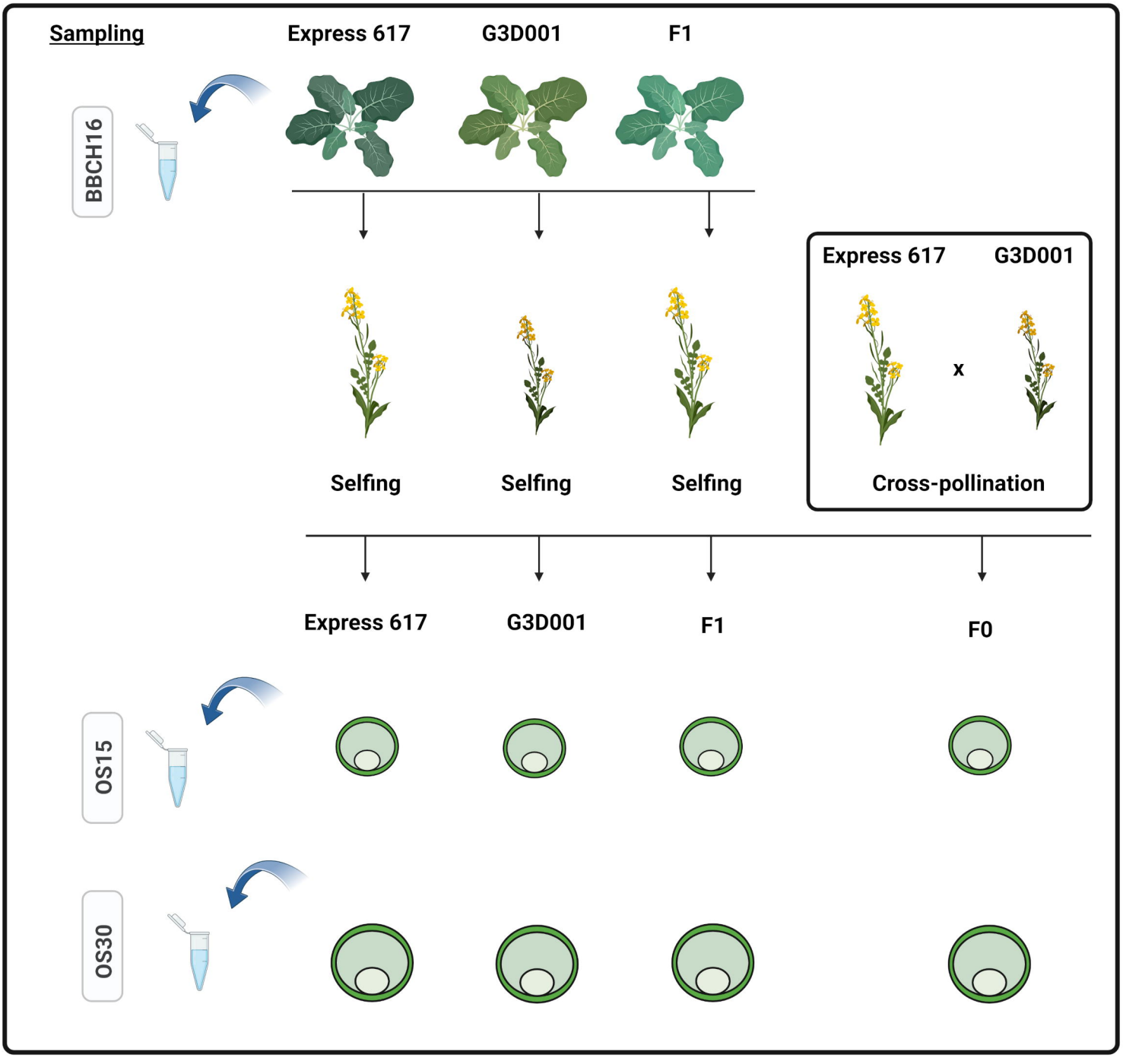
Zhao, D., Hamilton, J. P., Vaillancourt, B., Zhang, W., Eizenga, G. C., Cui, Y., et al. (2018). The unique epigenetic features of Pack-MULEs and their impact on chromosomal base composition and expression spectrum. *Nucleic Acids Res* 46, 2380–2397. doi: 10.1093/nar/gky025

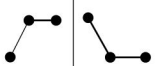
Zhu, A., Wang, A., Zhang, Y., Dennis, E. S., Peacock, W. J., and Greaves, A. I. K. (2020). Early Establishment of Photosynthesis and Auxin Biosynthesis Plays a Key Role in Early Biomass Heterosis in Brassica napus (Canola) Hybrids. *Plant Cell Physiol* 61, 1134–1143. doi: 10.1093/pcp/pcaa038

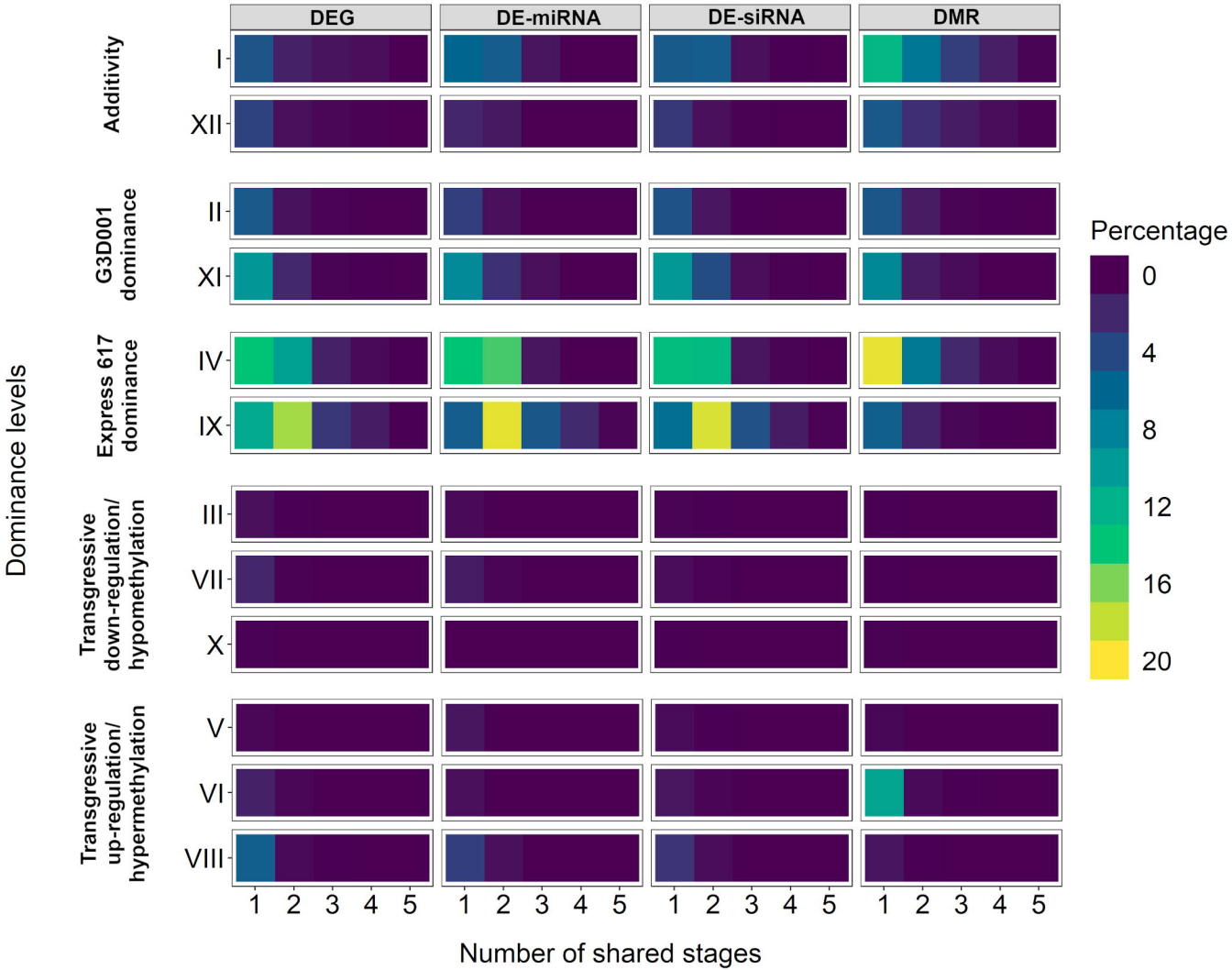
Ziegler, D. J., Khan, D., Kalichuk, J. L., Becker, M. G., and Belmonte, M. F. (2019). Transcriptome landscape of the early Brassica napus seed. *J Integr Plant Biol* 61, 639–650. doi: 10.1111/jipb.12812

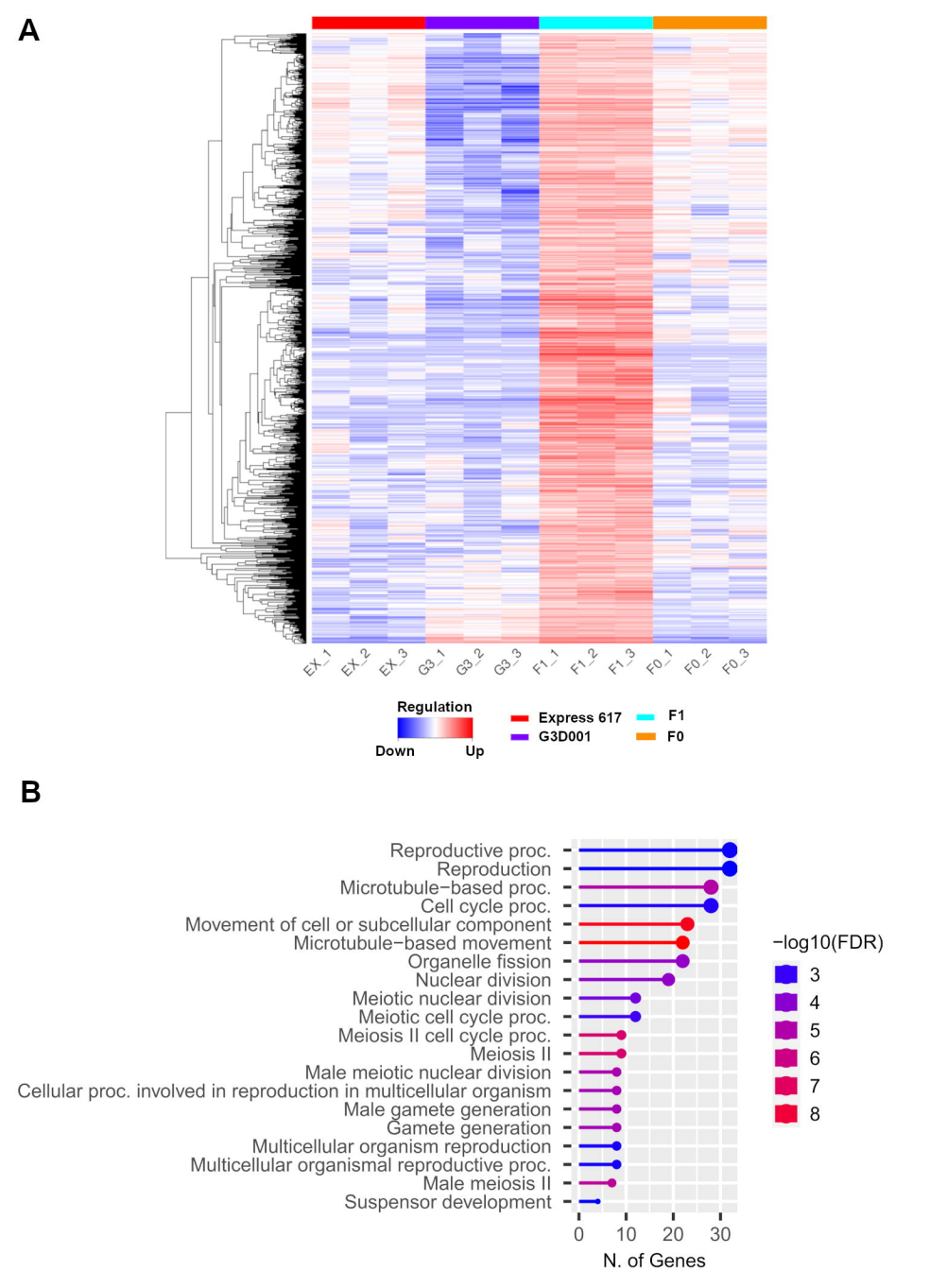
Zou, J., Hu, D., Mason, A. S., Shen, X., Wang, X., Wang, N., et al. (2018). Genetic changes in a novel breeding population of Brassica napus synthesized from hundreds of crosses between B. rapa and B. carinata. *Plant Biotechnol J* 16, 507–519. doi: 10.1111/pbi.12791

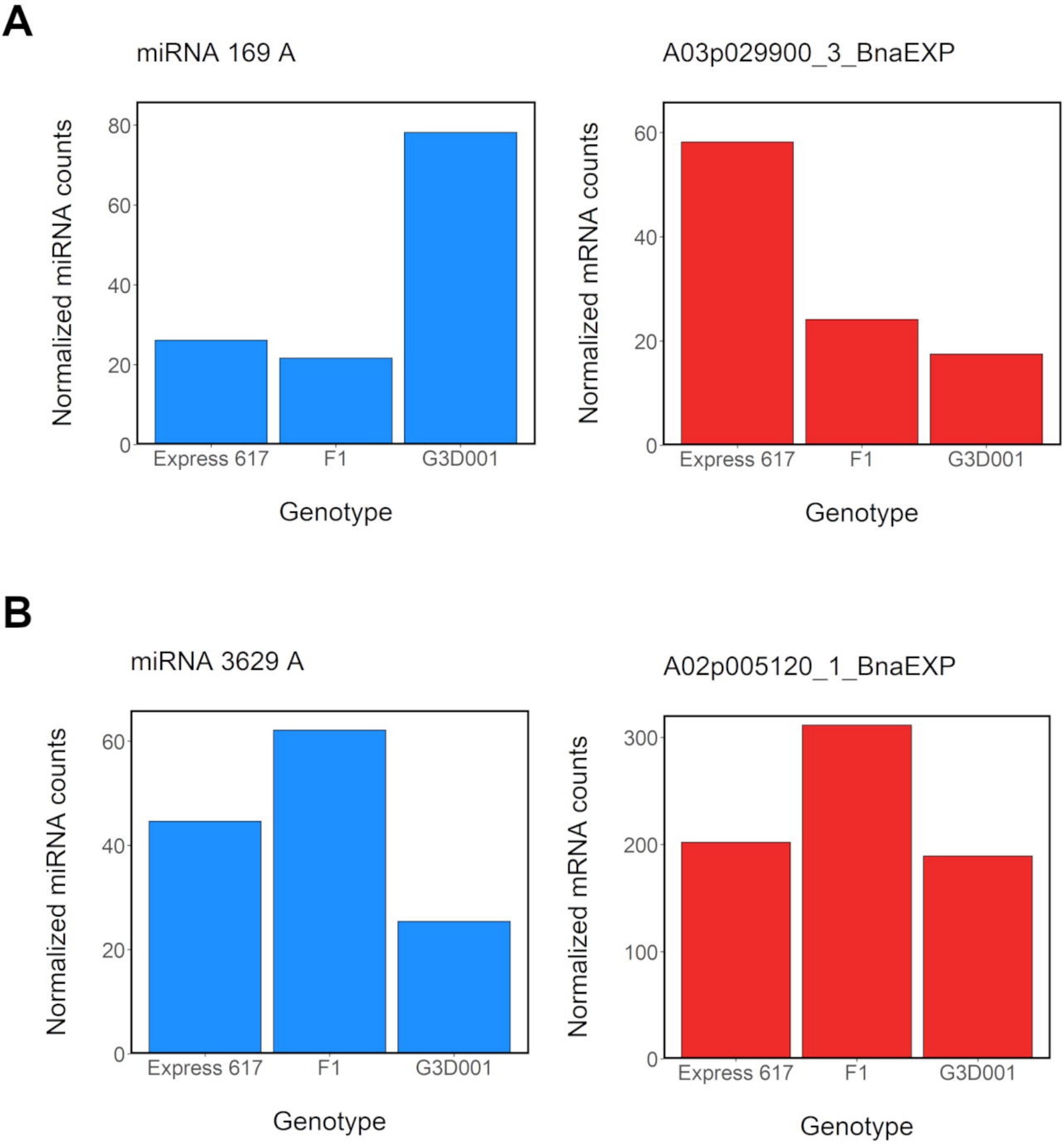
Zrimec, J., Börlin, C. S., Buric, F., Muhammad, A. S., Chen, R., Siewers, V., et al. (2020). Deep learning suggests that gene expression is encoded in all parts of a co-evolving interacting gene regulatory structure. *Nat Commun* 11, 6141. doi: 10.1038/s41467-020-19921-4

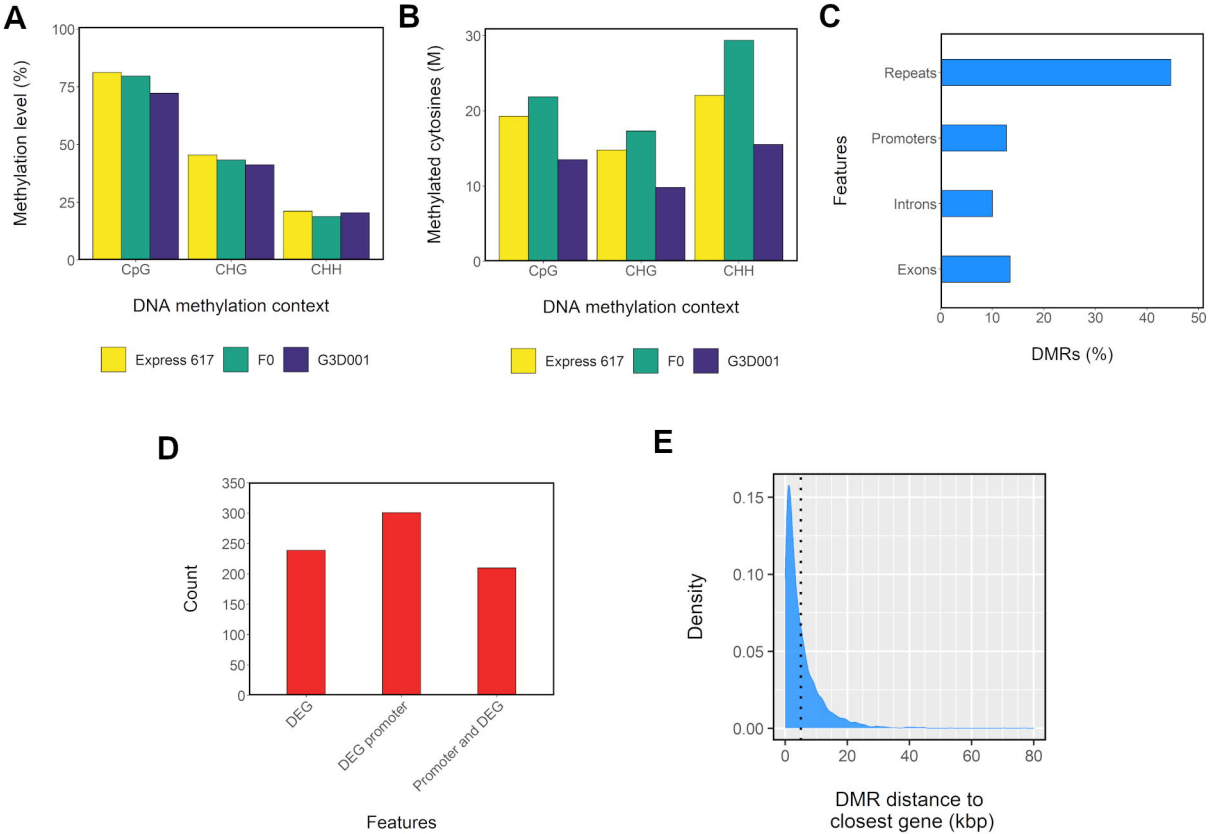

Sampling
Express 617
G3D001
F1
BBCH16
Selfing
Selfing
Selfing
Express 617
G3D001
x
Cross-pollination
Express 617
G3D001
F1
F0
OS15
OS30

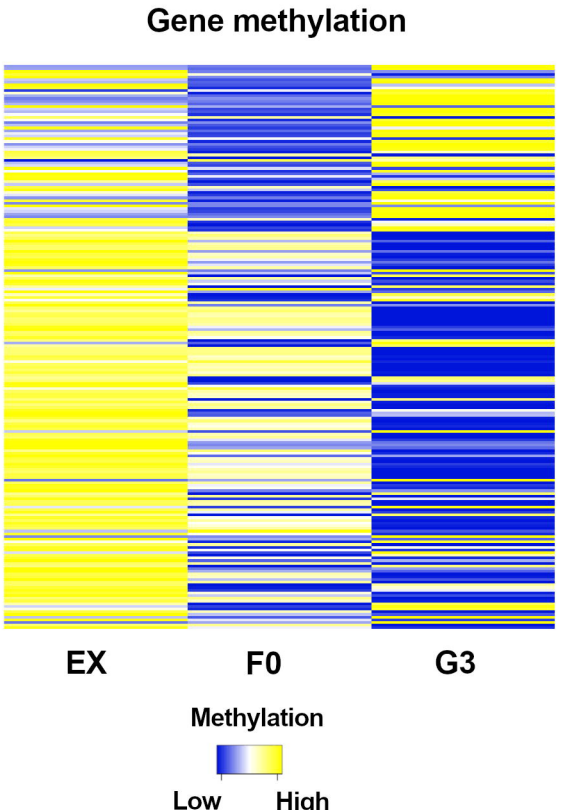
Feature	Stage	Additivity		G3D001 dominance		Express 617 dominance		Transgressive-down regulation/hypomethylation			Transgressive up regulation/hypermethylation		
		I	XII	II	XI	IV	IX	III	VII	X	V	VI	VIII
													
		E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G	E-F-G
DEG	BBCH16	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	0.00	0.49	0.52	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

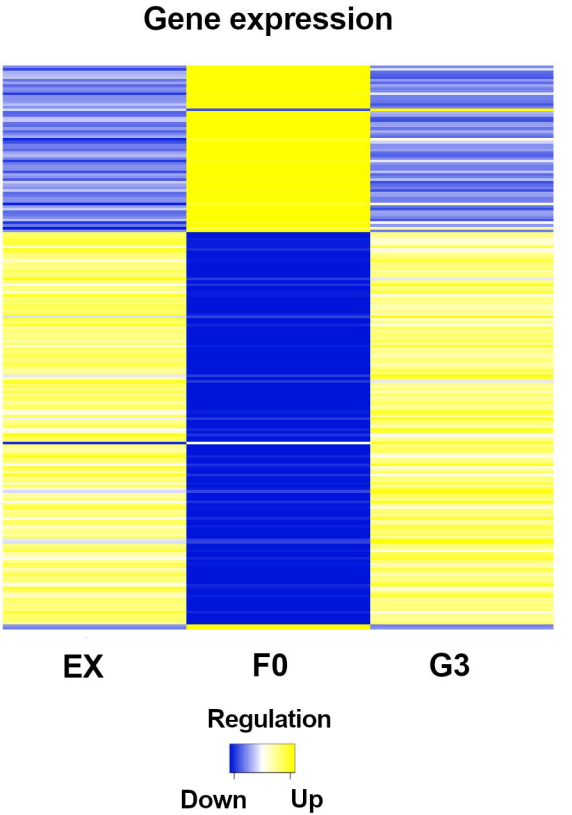

DEG
DE-miRNA
DE-siRNA
DMR
Additivity
I
XII
G3D001 dominance
II
XI
Express 617 dominance
IV
IX
Transgressive down-regulation/ hypomethylation
III
VII
X
Transgressive up-regulation/ hypermethylation
V
VI
VIII
Dominance levels
1 2 3 4 5
Number of shared stages
Percentage
0
4
8
12
16
20

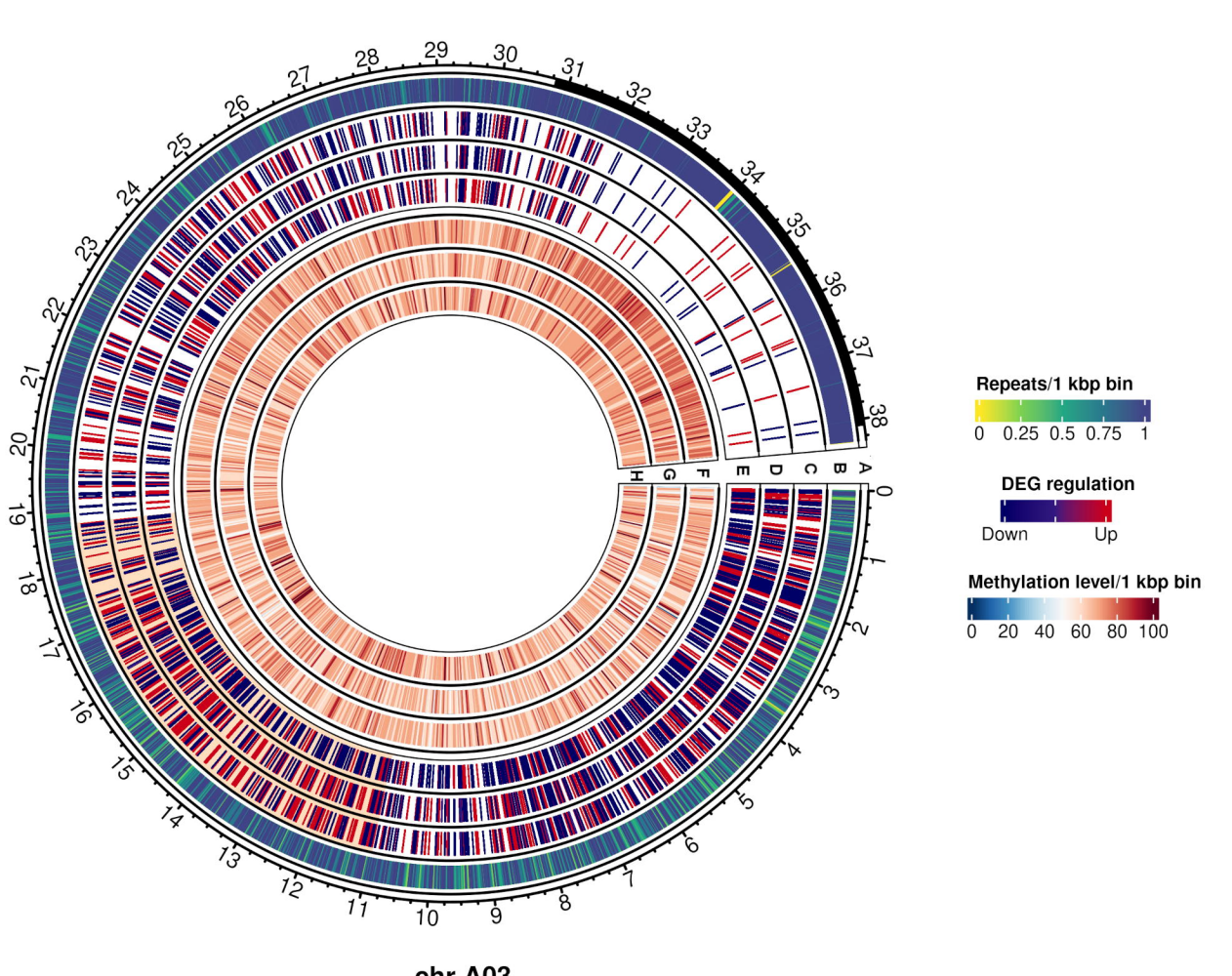

A
EX_1
EX_2
EX_3
G3_1
G3_2
G3_3
F1_1
F1_2
F1_3
F0_1
F0_2
F0_3
Regulation
Down
Up
Express 617
G3D001
F1
F0
B
Reproductive proc.
Reproduction
Microtubule−based proc.
Cell cycle proc.
Movement of cell or subcellular component
Microtubule−based movement
Organelle fission
Nuclear division
Meiotic nuclear division
Meiotic cell cycle proc.
Meiosis II cell cycle proc.
Meiosis II
Male meiotic nuclear division
Cellular proc. involved in reproduction in multicellular organism
Male gamete generation
Gamete generation
Multicellular organism reproduction
Multicellular organismal reproductive proc.
Male meiosis II
Suspensor development
0
10
20
30
N. of Genes
−log10(FDR)
3
4
5
6
7
8


A
miRNA 169 A
Normalized miRNA counts
0
20
40
60
80
Express 617
F1
G3D001
Genotype
A03p029900_3_BnaEXP
Normalized mRNA counts
0
20
40
60
Express 617
F1
G3D001
Genotype
B
miRNA 3629 A
Normalized miRNA counts
0
20
40
60
Express 617
F1
G3D001
Genotype
A02p005120_1_BnaEXP
Normalized mRNA counts
0
100
200
300
Express 617
F1
G3D001
Genotype


A
Methylation level (%)
100
75
50
25
0
CpG
CHG
CHH
DNA methylation context
Express 617
F0
G3D001
B
Methylated cytosines (M)
30
20
10
0
CpG
CHG
CHH
DNA methylation context
Express 617
F0
G3D001
C
Features
Repeats
Promoters
Introns
Exons
0
10
20
30
40
50
DMRs (%)
D
Count
350
300
250
200
150
100
50
0
DEG
DEG promoter
Promoter and DEG
Features
E
Density
0.15
0.10
0.05
0.00
0
20
40
60
80
DMR distance to
closest gene (kbp)


Gene methylation
EX
F0
G3
Methylation
Low
High


Gene expression
EX
F0
G3
Regulation
Down
Up


Repeats/1 kbp bin
0 0.25 0.5 0.75 1
DEG regulation
Down
Up
Methylation level/1 kbp bin
0 20 40 60 80 100
A
B
C
D
E
F
G
H
chr A03