Current Plant Biology 35-36 (2023) 100306

Contents lists available at ScienceDirect

Current Plant Biology

journal homepage: www.elsevier.com/locate/cpb

Integrated analysis of transcriptomic and small RNA sequencing data provides miRNA candidates for engineering agronomically important seed traits in *Brassica juncea*

Rubi Jain [a], Namrata Dhaka [b,*], Pinky Yadav [b], Manoj Kumar Sharma [c], Md Danish [b,1], Shalu Sharma [b,1], Sonika Kumari [b,1], Ira Vashisht [c], RK Brojen Singh [a], Rita Sharma [d]

[a] School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi, India
[b] Department of Biotechnology, School of Interdisciplinary and Applied Sciences, Central University of Haryana, Mahendergarh, Haryana, India
[c] School of Biotechnology, Jawaharlal Nehru University, New Delhi, India
[d] Department of Biological Sciences, Birla Institute of Technology and Science (BITS) Pilani, Pilani Campus, Rajasthan, India

ARTICLE INFO

Keywords:
Brassica
MicroRNA
Oil content
Seed coat color
Seed development
Seed size

ABSTRACT

Brassica juncea L. is an important oilseed crop that yields edible oil and biofuel. Improving *B. juncea* seed traits is a primary breeding target, but these traits are genetically complex. MicroRNAs (miRNAs) regulate seed development by modulating gene expression at the post-transcriptional or translational level and are excellent candidates for improving seed traits. However, the roles of miRNAs in *B. juncea* seed development are yet to be investigated. Here, we report small RNA profiling and miRNA identification from developing seeds of two contrasting varieties of *B. juncea*, Early Heera2 (EH2) and Pusa Jaikisan (PJK). We identified 326 miRNAs, including 127 known and 199 novel miRNAs, of which 103 exhibited inter-varietal differential expression. Integrating miRNAome and our previous transcriptome data identified 13,683 putative miRNA-target modules. Segregation of differentially expressed miRNAs into different groups based on variety-wise upregulation, followed by comprehensive functional analysis of targets using pathway mapping, gene ontology, transcription factor mapping, and candidate gene analysis, revealed at least 11, 6, and 7 miRNAs as robust candidates for the regulation of seed size, seed coat color, and oil content, respectively. Further, co-localization with previously reported quantitative trait loci (QTL) proffered 29 and 15 miRNAs overlapping with seed weight and oil content QTLs, respectively. Our study is the first comprehensive report of miRNAome expression dynamics from developing seeds and provides candidate miRNAs and target genes for engineering seed traits in *B. juncea*.

1. Introduction

Brassica juncea (Indian mustard) is a vital oilseed crop of the Indian subcontinent and various other regions of the world [20]. It is an allotetraploid (AABB, 2n = 36) species that evolved after multiple hybridization events between *B. rapa* (AA) and *B. nigra* (BB). Since it is an important source of edible oil and biofuels, boosting the productivity and quality of seed oil is a major breeding goal. Therefore, seed traits like seed weight, oil content, seed coat color, etc., are principal improvement targets. These traits are governed by many genes, signaling molecules, phytohormones, and noncoding RNAs that form intricate networks [14].

MicroRNAs (miRNAs) are a class of noncoding RNAs with a length of 20–24 nucleotides (nt). They are transcribed from MIR genes as primary miRNAs (pri-miRNAs) by DNA-dependent RNA polymerase II (Pol II) activity [71]. These pri-miRNAs are single-stranded and polyadenylated. They are folded into stem-loop hairpin-like structures and undergo processing within the Dicing bodies (D-bodies). Thereafter, several proteins such as HYPONASTIC LEAVES 1 (HYL1), zinc-finger protein SERRATE (SE), and TOUGH (TGH) recognize pri-miRNAs which are then cleaved by DICER-LIKE1 (DCL1) enzyme to release mature miRNA/miRNA* duplexes with 2-nt overhangs at 3′ ends. Subsequently, the miRNA duplex is methylated by HUA ENHANCER 1 (HEN1) on its 2′-O-methylation at 3′end [73]. The miRNA/miRNA*

* Correspondence to: Room no. 17, Department of Biotechnology, Acadmic Block 1, Central University of Haryana, Mahendergarh 123031, Haryana, India.
E-mail addresses: namratadhaka@gmail.com, namrata@cuh.ac.in (N. Dhaka).
[1] Equal authors

https://doi.org/10.1016/j.cpb.2023.100306
Received 6 September 2023; Received in revised form 14 November 2023; Accepted 22 November 2023
Available online 24 November 2023

duplexes are then exported to the cytoplasm, where miRNA is loaded on an Argonaute 1 (AGO1) protein to form an miRNA-inducing slicing complex (miRISC), while miRNA* is usually degraded [9].

More than 20 miRNAs have been experimentally validated in regulating seed development, and high throughput studies indicate the involvement of many more in modulating seed traits [14]. For instance, miR167, miR396, and miR408 regulate seed size and/or seed weight in several species. MiRNA biogenesis machinery genes like *DCL1, AGO1, HEN1*, etc., are also critical for embryogenesis during seed development, and their mutants exhibit severe abnormalities related to embryo development. Further, miRNAs like miR156, miR166, and miR167 are components of the *LEC1, ABI3, FUS3, LEC2* (LAFL) network, which is at the core of a typical dicot seed development program. These together regulate seed maturation as well as the accumulation of storage products.

Among the oilseed Brassicas, various studies have utilized small RNA profiling to dissect seed traits using developing seeds of *B. napus* [19,24,58,64,66,67,78]. However, miRNA dynamics during seed development have not been analyzed yet from Indian mustard, *B. juncea*.

Indian mustard has previously been genetically characterized to possess two distinct gene pools offering genetically unique resources for heterotic breeding [57]. *B. juncea* varieties Early Heera2 (EH2) and Pusa Jaikisan (PJK) belong to East-European and Indian origin, respectively, and exhibit marked differences in seed size, oil content, and seed coat color [10,50]. EH2 is a yellow and small-seeded variety with an oil content of ~36.9% and a thousand-seed weight of 2.0 g. On the other hand, PJK is a brown and bold-seeded variety with an oil content of ~41.1% and a thousand seed weight of 6.9 g [10,50]. The genetic distinctness of these varieties has also been leveraged to develop a double haploid mapping population called EPJ (EH2 x Pusa Jaikisan) [13]. Genetic analyses of this population have resulted in the identification of QTLs regulating seed weight and oil content [13,50]. Besides, transcriptome analysis of developing seeds from EH2 and PJK has previously uncovered vital genes and pathways responsible for varietal differences in seed size [10,36]. In our earlier transcriptome analysis, we ascertained that developing seeds at 15 and 45 days after pollination were most informative in discovering key genes associated with inter-varietal differences [10]. Here, we report the identification and analysis of miRNAs from seeds collected at these two stages from EH2 and PJK varieties. Further, we integrated small RNA findings with transcriptome data to shortlist candidates responsible for determining seed size, seed coat color and oil content.

2. Materials and methods

2.1. Growing of plants and harvesting the developing seeds

B. juncea varieties EH2 and PJK were grown from October to March 2021 under open field conditions at Jawaharlal Nehru University, New Delhi, India, and developing seeds were harvested at two different stages, 15 and 45 days after pollination (D), corresponding to early and late stages of seed development as demarcated in our previous study [10]. The seeds were immediately frozen in liquid nitrogen and stored at − 80 °C.

2.2. RNA isolation and small RNA sequencing

Total RNA was isolated using TRIzol™ (Ambion) reagent, described in Dhaka et al., [11], with minor modifications. To eliminate genomic DNA from RNA samples, total RNA was treated with DNase-I using TURBO DNA-free™ Kit (AM1907, Invitrogen) as per the manufacturer's instructions. The quantity and quality of RNA were checked using the Nanodrop Spectrophotometer, agarose gel electrophoresis, and Agilent 2200 Tape Station. RNA samples with RIN values > 6 were used for small RNA library preparation using QIAseq® miRNA Library Kit (Qiagen, Maryland, U.S.A.). Single-end sequencing was performed using Illumina NovaSeq 6000 sequencing machine with an average read length of 50 bp.

2.3. Analysis of small RNA sequencing data

The sequencing raw data was obtained in FASTQ files, and the data quality was assessed by FASTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapter sequences were trimmed, and reads ranging from 16 to 40 nucleotides (nt) in length were retained using Cutadapt v.4.1 [35] with parameters -a AACTGTAGGCACCATCAAT -m 16 -M 40. High-quality reads with a Phred score $\leq$ of 30 for a minimum of 70% of bases were retained using the FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/commandline.html). From the reads thus filtered, unique reads and their counts were extracted from each sample by collapsing the reads using fastx_collapser from FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/commandline.html). Redundant reads were removed for further processing, and the FASTQ format of unique reads was converted to FASTA format. Subsequently, unique reads were used to identify miRNAs in each sample.

2.4. Identification of known and novel miRNAs

To identify miRNAs, we selected sequences ranging from 20 to 24 nt and used the miRDeep-P2 program, which identifies miRNAs based on its plant-specific scoring system [26]. We first screened the unique reads against the Rfam database (v.14.9) [22] for removal of housekeeping non-coding RNAs, including ribosomal RNAs (rRNAs), transfer RNAs (tRNA), small nuclear RNAs (snRNA), and small nucleolar RNAs (snoRNAs). Next, the small RNA reads were mapped to the reference genome of the *B. juncea* variety Varuna [46] and potential miRNA precursors of the optimal size were extracted. Subsequently, the RNA secondary structures of these possible miRNA precursors were predicted, and the minimum free energy (MFE) was calculated. The criteria used for miRNA identification were as follows: a read per million (RPM) threshold of 1, zero mismatch, a maximum precursor length of 300 nt, allowing one thread for RNAfold, and a maximum of 15 different locations for read mapping related to miRNAs [26]. The predicted miRNAs from each sample were aligned against the miRBase database v.22.1 (http://www.mirbase.org/) [25] using Bowtie (v.1.3.2) [27], allowing one mismatch. The miRNAs that mapped to miRBase were identified as known miRNAs, while the remaining were novel miRNAs. The properties of miRNA precursors were assessed using the minimum free energy (MFE), and the values were normalized by the minimum free energy index (MFEI). The MFEI was calculated using the formula: MFEI = (MFE/(length of miRNA precursor sequence)* 100/(G+C) %), where MFE represents the negative folding free energy of miRNA secondary structures [72]. Further, the genomic location of miRNA precursor sequences was determined using bedtools (https://bedtools.readthedocs.io/en/latest/) and the sequence coordinates were illustrated on the *B. juncea* chromosomes using Mapchart v.2.32 [63].

2.5. Expression analysis of miRNAs

The number of reads representing unique miRNAs was obtained from the sequencing data for each sample. The miRNA expression levels were normalized in terms of reads per million (RPM) using the formula: normalized expression = (miRNA read count/total count of clean reads) * 10^6. Pairwise differential expression analysis of miRNAs was performed using DESeq2 v.1.38.2 [33]. The miRNAs showing absolute log2 fold change ≥ 1 or ≤ -1 and Benjamini Hochberg correction with FDR (false discovery rate) ≤ 0.05 were considered as differentially expressed. Heatmaps for expression visualization were prepared using MultiExperiment Viewer (MeV) [18].

2.6. Identification and functional analysis of miRNA targets

For the prediction of miRNA targets, the in-house generated transcriptomic dataset from EH2 and PJK seeds at different stages of development (15D, 30D, and 45D), comprising a total of 112,550 transcripts, was used (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA824648) [10]. Putative targets were predicted using the psRNA-Target web server [8] with an expectation value cut-off of 3 and a maximum unpair energy (UPE) of 25. The miRNA-target network was analyzed using Cytoscape software v.3.10.1 [53].

The *Arabidopsis* orthologs of target genes were retrieved using BLASTx with TAIR 10 proteins (https://www.arabidopsis.org/download/index-auto.jsp?dir=%2Fdownload_files%2FProteins). Pathway mapping for the target genes was done using MapMan (https://mapman.gabipd.org/mapman). Transcription factor mapping was done using Plant Transcription Factor Database (PTFDB v5.0) (http://planttfdb.gao-lab.org/). Enrichment analyses for pathway and transcription factor (TF) enrichment in each category were carried out using a hypergeometric test by Benjamini Hochberg correction with an FDR of ≤ 0.05. GO enrichment analysis of differentially expressed miRNA-target genes was performed by AgriGO (http://systemsbiology.cau.edu.cn/agriGOv2/c_SEA.php) using a p-value cut-off of ≤ 0.05 followed by a hypergeometric test with FDR of ≤ 0.05. Candidate gene search was carried out using a comprehensive literature survey as well as different databases like ARALIP (http://aralip.plantbiology.msu.edu/pathways/pathways), BRAD (http://brassicadb.cn/#/), TAIR (https://www.arabidopsis.org/), and SEEDGENES (https://seedgenes.org/).

2.7. Identification of miRNAs and targets co-localizing with the QTLs associated with seed weight and oil content

Seed weight and oil content QTL information for the EPJ population was downloaded from previous studies of *B. juncea* [13,50]. The information on the EPJ map was downloaded from Dhaka et al., [12]. The primer sequence information from [12] was used to determine the physical coordinates of the flanking markers of each QTL using BLASTn [1] against the *B. juncea* genome [46]. The QTLs for which physical coordinates could not be determined unambiguously were removed from further analysis. The QTLs were depicted on respective chromosomes using Mapchart v.2.32 [63].

2.8 miRNA-target expression analysis by quantitative real-time PCR (qRT-PCR).

miRNAs were reverse transcribed using Mir-X™ miRNA First-Strand Synthesis Kit (Takara). qRT-PCR reactions were performed using cDNA and miRNA-specific forward and universal reverse primers using TB Green Advantage® qPCR Premix (Takara), following the manufacturer's instructions. The reaction was performed using the following PCR cycle: initial denaturation step at 95 °C for 10 s, followed by 40 cycles of denaturation at 95 °C for 5 s, and annealing/extension at 55 °C for 20 s. First-strand cDNA synthesis of target genes was performed using iScriptTM cDNA Synthesis Kit (BioRad), and qPCR was carried out using SYBR™ Green PCR Master Mix (Applied Biosystems). The reaction conditions were as follows: initial denaturation at 95 °C for 10 min, followed by 40 cycles of denaturation for 15 s at 95 °C, annealing for 30 s at 55–60 °C and extension for 30 s at 72 °C. All qRT-PCR analyses were performed in Stratagene™ Mx3005P (Agilent Technologies) according to the manufacturer's instructions. Three technical replicates for each of the three biological replicates were analyzed. The *U6* snRNA (as provided in the kit) and *TIPS41* [5] were used as internal controls for miRNAs and targets, respectively. The relative expression levels of miRNAs and their respective target genes were evaluated following the delta-delta C_T method [32]. The primers used for the analyses are listed in Supplementary Table 17.

3. Results

3.1. Small RNA sequencing from early and late seed development stages

We performed small RNA sequencing of twelve libraries that comprise three biological replicates of early (15D) and late (45D) stages of developing seeds from *B. juncea* varieties, EH2 and PJK. More than 123 million raw reads were generated after sRNA profiling, with ~ 57 million and ~66 million reads obtained from EH2 and PJK, respectively. After adapter trimming, length filtering (16–40 nt), and removing the low-quality reads, ~32 and ~46 million high-quality reads were retained, corresponding to ~7 and ~13 million unique reads in EH2 and PJK, respectively (Fig. 1a, Supplementary Table 1). The size distribution analysis showed that the maximum number of reads were 24 nt in length in all the samples (Fig. S1). Pearson's correlation coefficient among the biological replicates of seed samples ranged from 0.96 to 1, except for two samples (E15-C and E45-B) exhibiting low correlation, which were removed from further analysis (Supplementary Table 2).

3.2. Identification and characterization of seed miRNAome of B. juncea

After removing the reads mapping to housekeeping noncoding RNAs such as rRNAs, tRNAs, snRNAs, and snoRNAs, miRNAs were identified using a conventional pipeline (Fig. 1a). A total of 326 miRNAs were identified. Subsequently, based on the mapping against the Viridiplantae mature miRNA sequences from the miRBase (Release 22.1), 127 miRNAs were classified as known and199 as novel (Fig. 1a, Supplementary Table 3). These 326 miRNAs mapped to 744 precursors, with 461 and 283 precursors for known and novel miRNAs, respectively (Supplementary Table 3).

The known miRNAs had precursor lengths ranging from 49 to 219 nt, with an average length of 99 nt, while the precursor lengths of the novel miRNAs varied from 42 to 227 nt, with an average size of 114 nt (Supplementary Table 4). Variety-wise analysis showed that 115 (35%) of the 326 miRNAs were detected in both EH2 and PJK and among these, 84 were known, and 31 were novel. On the other hand, 95 (29%) (12 known and 83 novel) were explicitly detected in EH2, while 116 (36%) (31 known and 85 novel) were PJK-specific (Fig. 1b). Furthermore, the stage-wise analysis for each variety showed that 66 (20%) miRNAs were detected in all the stages of both varieties (Fig. 1c), including 55 (17%) known and 11 (3%) novel miRNAs.

For nomenclature, each unique mature miRNA sequence was named using the prefix 'Bju', followed by the miRNA family name and number. Different mature miRNA sequences belonging to the same miRNA family were named by adding a numeric suffix after the family name (such as Bju-miR156.1, Bju-miR156.2, and so on) (Supplementary Table 3). Similarly, the novel miRNAs were named from Bju-novel-miR1 to Bju-novel-miR214 (Supplementary Table 3). The precursors of known miRNAs were named Bju-MIR156.1–1, Bju-MIR156.1–2, Bju-MIR156.2, and so on, while novel miRNA precursors were named Bju-novel-MIR1 to Bju-novel-MIR214.

We first characterized the length and nucleotide distribution to assess the sequence features of the *B. juncea* seed miRNAome. The length distribution showed that 96 (76%) of the known miRNAs were 21 nt in length, whereas the remaining belonged to 20 nt and 22 nt categories (16 (13%) and 15 (12%), respectively) (Fig. 1d). The novel miRNAs also depicted a peak at 21 nt with the maximum number of 125 (63%) miRNAs, while 20 (10%), 49 (25%), 3 (1%), and 2 (1%) novel miRNAs belonged to 20, 22, 23, and 24 nt categories, respectively (Fig. 1e). The nucleotide distribution at their 5′ end showed that 67 (53%) of the known miRNAs contained uracil at the first position, followed by 33 (26%) with guanine, 15 (12%) with adenine, and 12 (9%) with cytosine (Fig. 1d). Similarly, among the novel miRNAs, 119 (60%) had uracil at their 5′ end, 29 (15%) had guanine, 26 (13%) had adenine, and 25 (12%) had cytosine (Fig. 1e).

The assignment of miRNA families placed 127 known miRNAs in 42

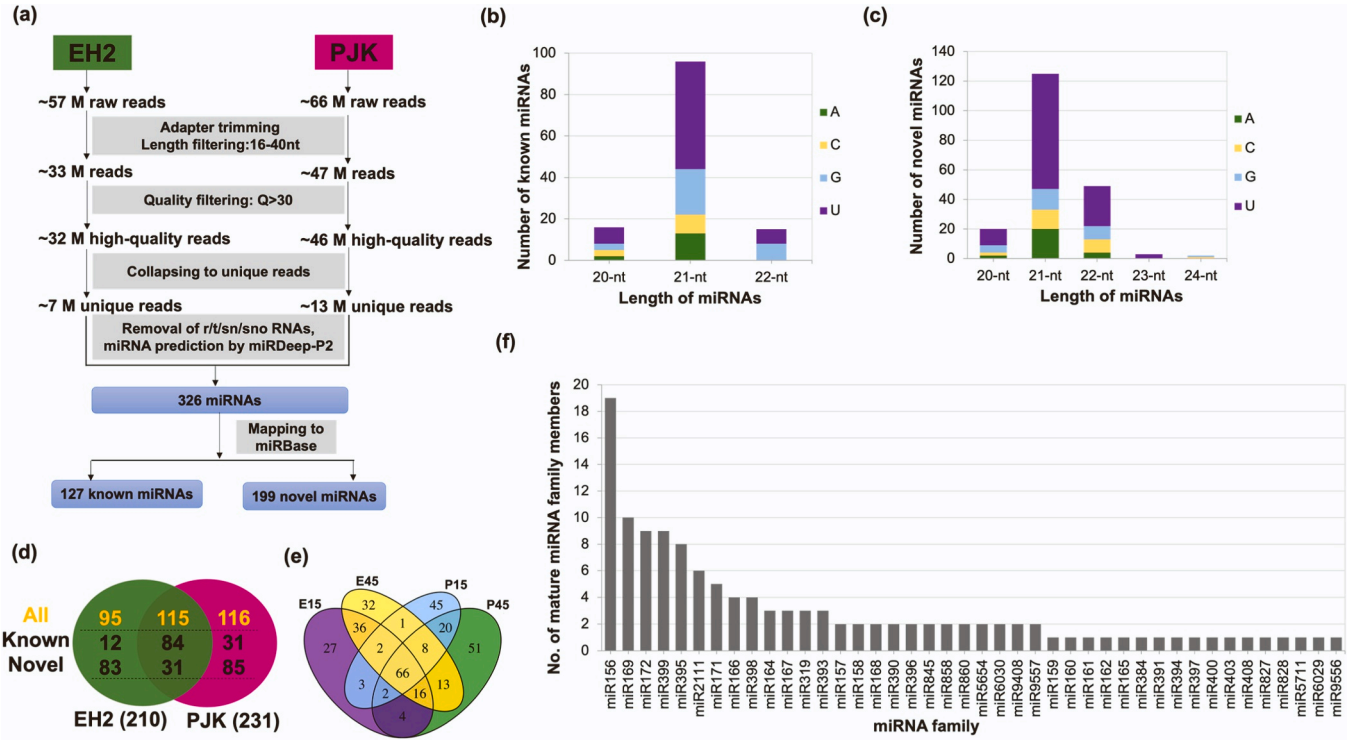

Fig. 1. Small RNA sequencing and identification of seed miRNAome from early and late seed development stages in *B. juncea*. (a) Workflow for identification of known and novel miRNAs from small RNA sequencing data obtained from developing seeds of Early Heera2 (EH2) and Pusa Jaikisan (PJK). M represents a million. (b) Size and 5′-end base distribution for known miRNAs. (c) Size and 5′-end base distribution for novel miRNAs. (d) Number of total known and novel miRNAs identified in EH2 and PJK. (e) Number of miRNAs identified at early (E15 and P15) and late stages (E45 and P45) in EH2 and PJK in stage-wise and variety-wise comparisons. (f) Number of mature miRNA family members identified from all the known miRNA families.

miRNA families (Fig. 1f). Among these, miR156 was the most abundant family with 19 members, followed by miR169 with ten members. Both miR172 and miR399 families comprised nine members each, while miR395 had eight members, the miR2111 family showed six members, and miR171 had five members. Two miRNA families (miR166 and miR398) comprised four members each, and four families (miR164, miR167, miR319, and miR393) comprised three members each. Moreover, 12 families (miR157, miR158, miR168, miR390, miR396, miR845, miR858, miR860, miR5654, miR6030, miR9408, and miR9557) had two members each, whereas 17 families (miR159, miR160, miR161, miR162, miR165, miR384, miR391, miR394, miR397, miR400, miR403, miR408, miR827, miR828, miR5711, miR6029, and miR9556) were represented by only a single member each (Fig. 1f).

The genomic localization showed that the 744 miRNA precursors were located on all 18 *B. juncea* chromosomes (Fig. S2). The chromosome B08 exhibited the highest number of miRNA precursors (60), while the chromosome A10 harbored 26 miRNA precursors. Further, A and B subgenomes harbored 366 and 378 miRNA precursors, respectively. We also observed clustering of miRNA precursors of the same family members in the genome (Fig. S2). Further, more than one cluster of the same miRNA family precursors was often located on the same or different *B. juncea* chromosomes. For instance, miR156 family members were on almost all chromosomes except for A05, B05, and B06. Members of miR169 and miR399 were distributed on 13 different chromosomes, while miR166, miR172, and miR319 were located on 12 different chromosomes.

The genomic coordinates showed that the majority, 564 (76%) of the miRNA precursors were located within the intergenic regions. In comparison, 80 (11%) were in exonic-intronic regions, 75 (10%) were derived from the exonic areas, and 25 (3%) from the intronic regions (Fig. S3a). The minimum free energy (MFE) calculation of the predicted secondary structures for miRNA precursors showed that most precursors had MFE values of − 60 to − 90 kcal/mol (Fig. S3b). The minimal folding energy index (MFEI) for all miRNA precursors was approximately 2 (Supplementary Table 4).

3.3. Expression dynamics of EH2 and PJK seed miRNAomes

The expression analysis was done using the reads per million data for all the miRNAs Supplementary Table 5). We observed that of the total 156 miRNAs expressed at the early stage in EH2, 8 (5%) exhibited low expression with RPM values ≤ 10, 57 (37%) showed expression levels ranging from 10 and 100 RPM, 44 (28%) had RPM values between 100 and 500, 11 (7%) exhibited expression levels between 500 and 1000, while 36 miRNAs (23%) showed RPM values > 1000 (Fig. 2a). We detected a similar trend in the early stage of PJK, with 5 out of 147 total miRNAs (3%) exhibiting low expression with RPM values ≤ 10, 37 miRNAs (25%) with expression levels between 10 and 100 RPM, 42 miRNAs (29%) showed RPM values between 100 and 500, 19 miRNAs (13%) had expression values between 500 and 1000, while 44 miRNAs (30%) exhibited RPM values > 1000 (Fig. 2a).

In the late stage, the expression patterns of miRNAs in EH2 revealed that of the total 174 miRNAs expressed in this stage, 16 (9%) showed expression with RPM values ≤ 10, 68 (39%) had expression levels ranging from 10 to 100 RPM, 37 (21%) showed expression levels between 100 and 500 RPM, 17 (10%) exhibited between 500 and 1000 RPM, and 36 (21%) showed RPM values > 1000 (Fig. 2a). Similarly, of the 180 miRNAs expressed in the late stage of PJK, 12 (7%) exhibited expression with RPM values ≤ 10, 67 (37%) showed expression levels between 10 and 100 RPM, 36 (20%) exhibited RPM values between 100 and 500, 21 (12%) exhibited RPM values between 500 and 1000, and 44 (24%) showed RPM values > 1000 (Fig. 2a).

Thus, at the early stage, the number of miRNAs expressed was slightly higher in EH2 than in PJK. However, the late stage showed more

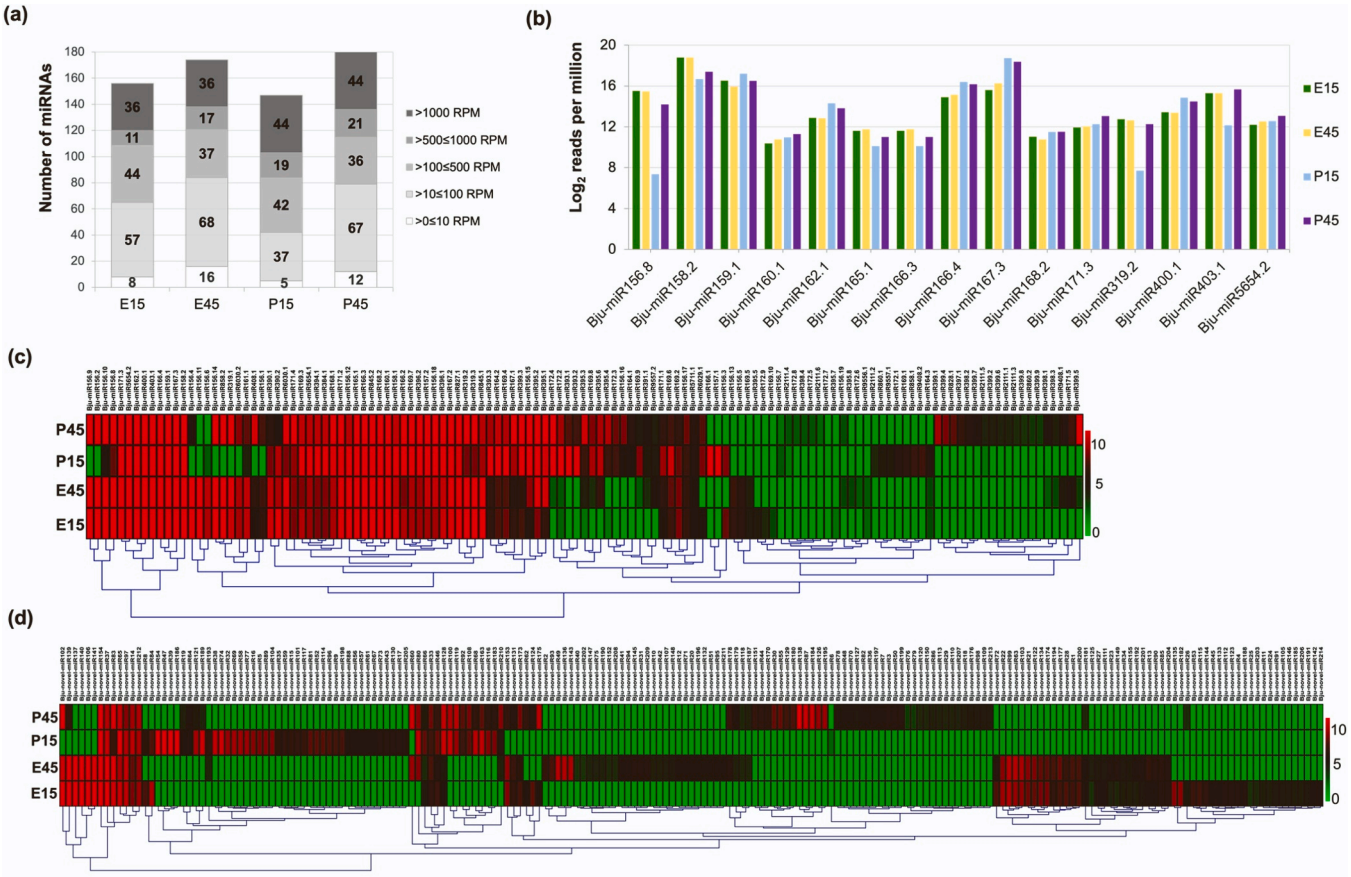

Fig. 2. Results of expression dynamics analysis of seed miRNAome. (a) Number of miRNAs expressed in EH2 (E) and PJK (P) at the early (15D) and late (45D) stages, according to ranges of RPM levels. (b) Expression levels of the top 15 known miRNAs transformed with $\log_2$ (RPM+1) at each stage, 15D and 45D. (c) Heatmap showing the expression of all the known miRNAs. (d) Heatmap showing expression of all the novel miRNAs.

miRNAs in PJK than EH2. Further, identification of highly expressed miRNAs in each of the stages in both the varieties showed that the top 15 miRNAs belonged to miR156, miR158, miR159, miR160, miR162, miR165, miR166, miR167, miR168, miR171, miR319, miR400, miR403, and miR5654 families (Fig. 2b). Interestingly, most of these miRNAs exhibited similar expression levels in all four categories, except miR156.8 and miR319.2, which showed low expression in P15 compared to the other stages.

We also identified the stage-specific and/or variety-specific miRNAs based on the expression levels. Of the total 326 miRNAs, 95 miRNAs (29%) were explicitly expressed in EH2 while 116 miRNAs (36%) were identified as PJK-specific (Supplementary Table 6). Within the EH2-specific miRNAs, 27 (28%) were expressed exclusively at the early stage, while 32 (34%) were late stage specific. Among the PJK-specific miRNAs, 45 (39%) were detected only in the early stage, while 51 (44%) were identified as late stage specific. Furthermore, we observed that 12 known miRNAs and one miRNA family, miR9556, were exclusively expressed in an EH2-specific manner, whereas 31 known miRNAs and four families, including miR397, miR828, miR860, and miR2111, were exclusively expressed in PJK-specific manner (Fig. 2c). Additionally, 83 novel miRNAs were exclusively expressed in an EH2-specific manner, while 85 other novel miRNAs were exclusive to PJK. In the stage-specific comparison, the miR398 family was exclusively expressed in the late stage (Fig. 2c, Supplementary Table 6). Bju-novel-miR8 and Bju-novel-miR84 were exclusively expressed in the early stage, while Bju-novel-miR60, Bju-novel-miR118, Bju-novel-miR178, Bju-novel-miR179, and Bju-novel-miR187 explicitly expressed in the late stage (Fig. 2d).

3.4. *Differentially expressed miRNAs*

The pairwise differential expression analysis between EH2 and PJK in both stages of seed development resulted in a total of 103 unique miRNAs, including 62 known and 41 novel miRNAs (Supplementary Table 7).

3.4.1. *Early-stage dynamics*

In the early stage, 51 known miRNAs were differentially expressed (DE), of which 25 and 26 were upregulated in EH2 and PJK, respectively (Fig. 3a). The 25 known miRNAs upregulated in EH2 belonged to 12 different families, including miR156, miR158, miR161, miR165, miR166, miR319, miR403, miR408, miR827, miR845, miR858, and miR6030. In PJK, the 26 known miRNAs belonged to 14 different families, such as miR156, miR157, miR158, miR164, miR166, miR167, miR169, miR172, miR384, miR393, miR394, miR395, miR396, and miR399 were upregulated. Also, different miRNA family members of miR156, miR158, and miR166 were upregulated in both varieties, while several miRNAs showed variety-specific upregulation. Some of these known miRNAs, like Bju-miR156.9, showed the highest differential accumulation in EH2 (14-fold), while Bju-miR172.4 had maximum upregulation in PJK (11-fold) (Supplementary Table 7).

In addition, 37 novel miRNAs were differentially expressed, of which 18 and 19 were upregulated in EH2 and PJK, respectively (Fig. 3b). Among them, Bju-novel-137 showed maximum upregulation (14-fold) in EH2, whereas Bju-novel-119 exhibited the maximum differential accumulation (13-fold) in PJK (Supplementary Table 7). The novel miRNAs like Bju-novel-16, Bju-novel-32, Bju-novel-39, Bju-novel-47, Bju-novel-54, Bju-novel-58, Bju-novel-69, Bju-novel-74, Bju-novel-77,

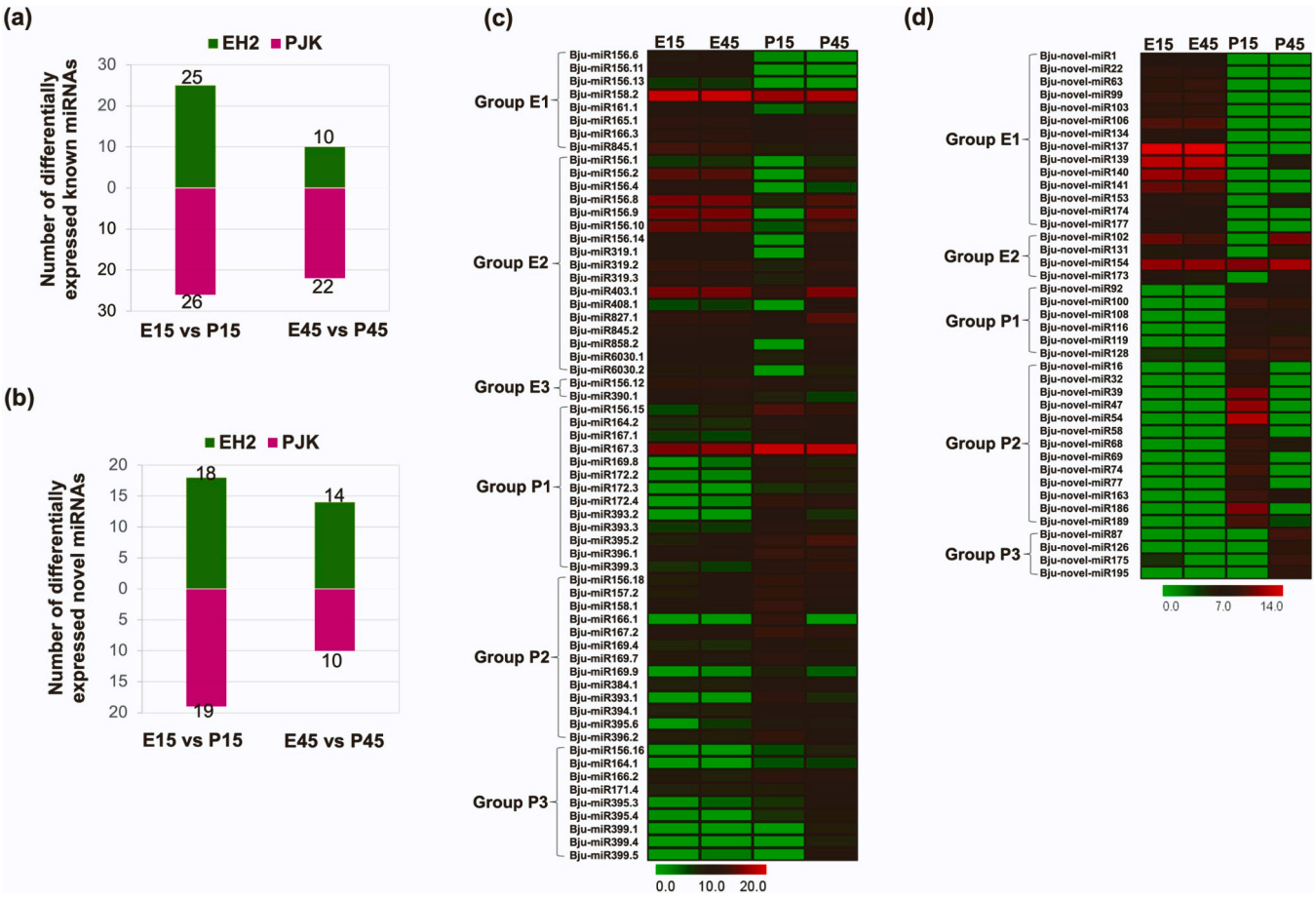

Fig. 3. Differentially expressed miRNAs. (a) Differentially expressed known miRNAs between varieties EH2 and PJK at early (15D) and late (45D) stages of seed development. (b) Differentially expressed known miRNAs between varieties EH2 and PJK at early (15D) and late (45D) stages of seed development. (c) Heatmap showing differentially expressed known miRNAs. Grouping was done based on variety-wise upregulation. Groups E1, E2, and E3 showed upregulation in EH2, while P1, P2, and P3 exhibited upregulation in PJK. (d) Heatmap showing differentially expressed novel miRNAs.

Bju-novel-miR92, Bju-novel-miR100, Bju-novel-miR108, Bju-novel-miR116, Bju-novel-miR119 and Bju-novel-186 were exclusively detected in PJK, while Bju-novel-miR1, Bju-novel-miR22, Bju-novel-miR63, Bju-novel-miR99, Bju-novel-miR103, Bju-novel-miR106, Bju-novel-miR134, Bju-novel-miR137, Bju-novel-miR140, Bju-novel-miR141, Bju-novel-miR174, and Bju-novel-miR177 were exclusive to EH2.

3.4.2. *Late-stage dynamics*

In the late stage, 32 known miRNAs were differentially expressed, of which 10 and 22 were upregulated in EH2 and PJK, respectively (Fig. 3a). In EH2, the known miRNAs belonged to families miR156, miR158, miR161, miR165, miR166, miR390, and miR845, whereas the known miRNAs upregulated in PJK belonged to families miR156, miR164, miR166, miR167, miR169, miR171, miR172, miR393, miR395, miR396, and miR399. This shows that various members of miRNA families of miR156 and miR166 were upregulated in both EH2 and PJK. Among the known miRNAs, Bju-miR156.11 showed the maximum differential accumulation in EH2 (9-fold), while Bju-miR172.4 exhibited maximum upregulation in PJK (31-fold) (Supplementary Table 7). Additionally, 24 novel miRNAs were differentially expressed in the late stage, of which 14 and 10 were upregulated in EH2 and PJK, respectively (Fig. 3b). Bju-novel-miR137 exhibited the highest differential accumulation in EH2 (16-fold), while Bju-novel-miR119 had maximum upregulation in PJK (8-fold) (Supplementary Table 7).

3.4.3. *Grouping of differentially expressed miRNAs based on variety-specific dynamics*

To further dissect the expression dynamics, we categorized the differentially expressed (DE) miRNAs into six groups based on the variety-specific differential accumulation (Supplementary Table 7). Groups E1, E2, and E3 exhibited EH2-preferential patterns while P1, P2, and P3 showed PJK-preferential expression patterns. Group E1 comprised miRNAs upregulated at both early and late stages in EH2 with 22 DE miRNAs (8 known and 14 novel) placed in this group (Fig. 3c & d). The known miRNAs in E1 belonged to families miR156, miR158, miR161, miR165, miR166, and miR845. Of these families, miR161 and miR165 were exclusive to this group. Group E2 contained 21 miRNAs (17 known and four novel) upregulated only in the early stage of EH2. The known miRNAs in this group belonged to families miR156, miR319, miR403, miR408, miR827, miR845, miR858, and miR6030. Interestingly, six of these families (miR319, miR403, miR408, miR827, miR858, and miR6030) were exclusive to this group (Fig. 3c & d). Similarly, group E3 consisted of miRNAs upregulated only in the late stage of EH2. This category harbored only two known miRNAs, Bju-miR156.12 and Bju-miR390.1, with miR390 exclusive to this group (Fig. 3c). Thus, overall, nine miRNA families miR161, miR165, miR319, miR390, miR403, miR408, miR827, miR858, and miR6030 were upregulated in EH2 only.

On the other hand, group P1 harbored miRNAs upregulated in PJK at both early and late stages (Fig. 3c & d). It included a total of 19 miRNAs (13 known and six novel miRNAs) belonging to miR156, miR164, miR167, miR169, miR172, miR393, miR395, miR396, and miR399

families. MiR172 was exclusive to this group. Group P2 comprised miRNAs upregulated in PJK in the early stage only and contained 26 DE miRNAs, with 13 known and 13 novel miRNAs. The known miRNAs belonged to families miR156, miR157, miR158, miR166, miR167, miR169, miR384, miR393, miR394, miR395, and miR396. MiR157 and miR394 were exclusive to this group. MiR167, miR169, and miR393, were exclusively present in groups P1 and P2 (Fig. 3c & d). Group P3 with miRNAs upregulated in PJK at a late stage only consisted of 13 miRNAs, including nine known and four novel miRNAs. The known miRNAs families were miR156, miR164, miR166, miR171, miR395, and miR399. MiR164 and miR399 were exclusive to groups P1 and P3, while miR395 was present in groups P1, P2, and P3. Thus, miRNAs from 9 families i.e., miR157, miR164, miR167, miR169, miR172, miR393, miR394, miR395, and miR399, were exclusively upregulated in PJK (Fig. 3c & d).

3.5. Identification of miRNA targets

To further elucidate the role of miRNAs during seed development in *B. juncea,* we used an in-house generated transcriptome dataset consisting of 112,550 transcripts expressed during *B. juncea* seed development [10] for target identification. Of these, 9796 unique transcripts were identified as putative targets of 326 miRNAs, resulting in 13683 miRNA-target modules (Supplementary Table 8). Of these, 12984 (94%) miRNA-target modules were predicted by psRNATarget to function through miRNA-directed target cleavage, while the remaining 870 (6%) modules likely function through translational inhibition.

Further, 7552 (55%) miRNA-target pairs showed an inverse correlation in their expression patterns in both varieties (Supplementary Table 8). The differentially expressed miRNAs and their targets formed 4804 miRNA-target modules (Supplementary Table 8). Further, a comparison of differentially expressed target transcripts using the previous transcriptome data of the corresponding stages [10] showed that a total of 457 miRNA-target transcripts pairs comprising 381 unique differentially expressed target transcripts and 93 unique differentially expressed miRNAs were inversely correlated.

3.6. Functional analysis of miRNA-target modules from each group

To further identify the miRNA candidates for roles in determining seed traits, we analyzed the potential functions of the miRNAs and their target transcripts. Ortholog identification showed 9796 unique target transcripts mapped to 4664 *Arabidopsis* genes (Supplementary Table 8). Moreover, 9234 unique target transcripts were assigned to 34 different MapMan pathways and 198 pathway sub-categories using the MapMan pathway mapping tool [62]. Additionally, 1012 unique target transcripts mapped to 47 different TF categories using the PTFDB [60] (Supplementary Table 8). Next, we analyzed the subset of 4804 pairs comprising differentially expressed miRNAs and respective targets and performed functional enrichment for pathways, GO, and TFs. We identified candidate miRNAs and target genes in each group (Supplementary Table 9).

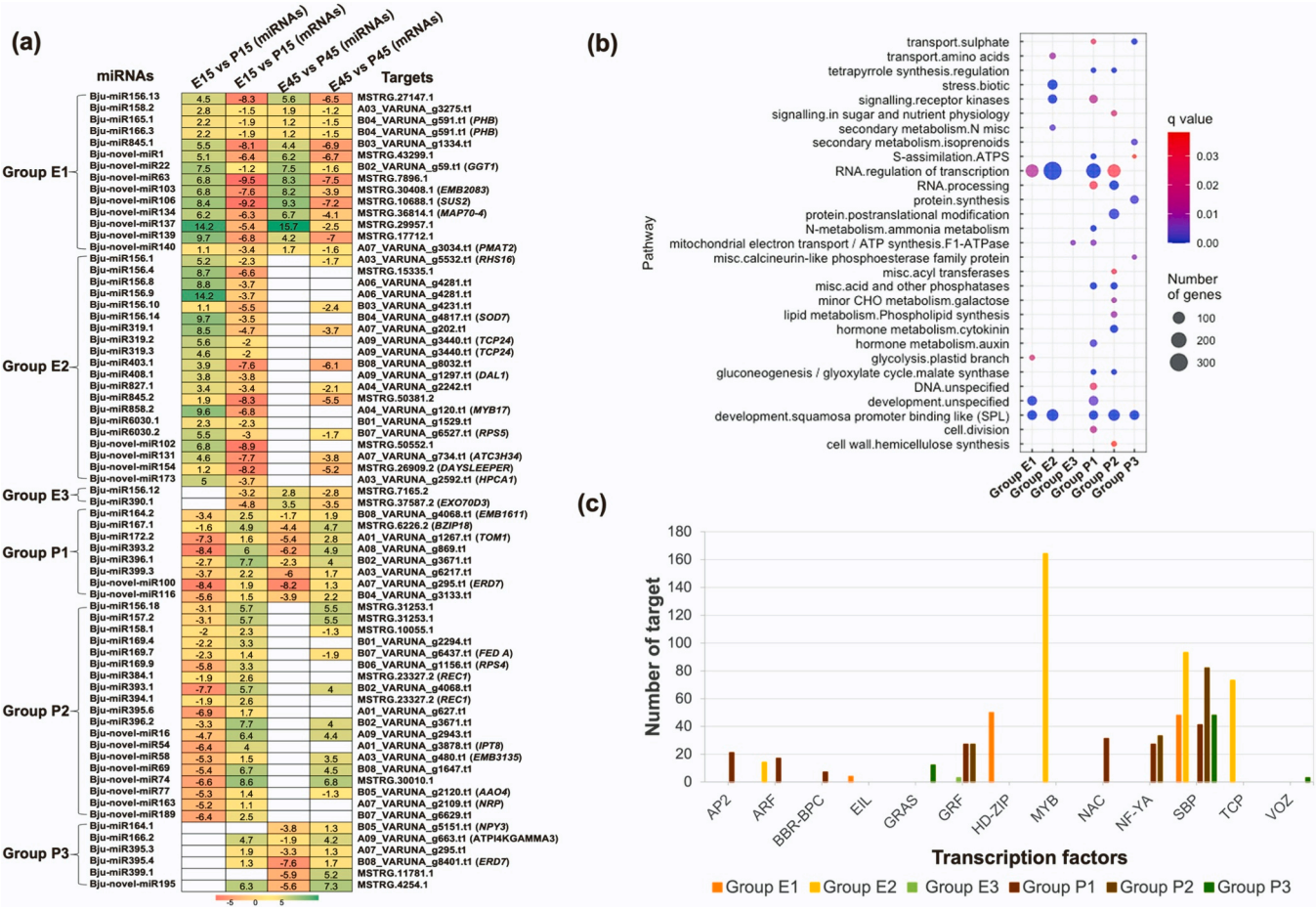

Fig. 4. Group-wise functional analysis of differentially expressed miRNAs and their target transcripts. (a) Heatmap showing differentially expressed miRNAs and respective differentially expressed target transcripts exhibiting an inverse correlation in each group. Due to multiple targets, only transcripts with the highest fold change are shown. (b) Results of pathway enrichment of target transcripts for each group. MapMan pathway sub-categories are shown on the left. The scale depicts the q value and the number of genes in each enriched category. (c) Results of transcription factor (TF) enrichment of target transcripts for each group. The bars show the number of enriched TFs in each group.

3.6.1. Group E1: miRNAs upregulated in both early and late stages of EH2

The target analysis showed 944 putative unique targets mapped to 22 DE miRNAs, comprising 990 miRNA-target modules (Supplementary Table 9). Of these, 752 miRNA-target transcript pairs showed an inverse correlation. Among these, 114 miRNA-target modules showed an inverse differential expression pattern with 14 miRNAs upregulated in both the early and late stages in EH2 while their corresponding 32 target transcripts were downregulated. The pairs with the highest fold change have been highlighted in Fig. 4a.

Pathway mapping revealed enrichment of 'development' and 'regulation of transcription' categories and TFs belonging to EIL, HD-ZIP, and SBP families (Fig. 4b & c, Supplementary Table 10 & 12). Further, candidate gene analysis based on the literature survey showed that the prominent known miRNAs of this group, miR156, miR165, and miR166, are involved in the regulation of cell division patterns and apical specifications during the seed development [37,42]. Further, mapping of the target transcripts with our previous transcriptome candidates [10] proffered several important miRNA-target modules like Bju-miR156.6-A07_VARUNA_g4239.t1, Bju-novel-miR1-A10_VARUNA_g839.t1, Bju-novel-miR63-A02_VARUNA_g4152.t1, Bju-novel-miR174-A06_VARUNA_g1175.t1/B05_VARUNA_g2542.t1, and Bju-novel-miR177 -B05_VARUNA_g4738.t1 as these correspond to transcripts which are important candidates for regulation of cell cycle and cell division, embryo development, and seed coat mucilage based on our previous transcriptome analysis [10].

3.6.2. Group E2: miRNAs upregulated only in the early stage of EH2

The target analysis showed 959 unique target transcripts for 21 miRNAs comprising 1162 miRNA-target modules (Supplementary Table 9). Of these, 782 miRNA-target transcript pairs showed an inverse correlation, and 145 pairs showed transcripts with inverse differential expression patterns. A total of 20 miRNAs were upregulated exclusively in the early stage, while corresponding 90 target transcripts were downregulated in the same stage.

Pathway mapping revealed enrichment of development, secondary metabolism, regulation of transcription, receptor kinases, and other categories (Fig. 4b, Supplementary Table 10). GO categories like cell differentiation, response to hormones like jasmonic acid, metabolic process, and response to lipids were enriched in this dataset (Supplementary Table 11). Further, TF enrichment showed that ARFs, MYBs, SBPs, and TCPs were overrepresented in this group (Fig. 4c, Supplementary Table 12).

Literature analysis revealed that miR156 and miR408 are crucial for cell division and seed size in the early stage of the EH2 [42,45]. Some of the crucial miRNA-target modules are Bju-miR156.1-A09_VARUNA_g2999.t1 and Bju-miR156.10-B01_VARUNA_g561.t1, Bju--miR156.14 -A09_VARUNA_g2999.t1/B04_VARUNA_g4817.t1, Bju-miR319.2/Bju-miR319.3 -MSTRG.35655.1, Bju-miR858.2-B01_VARUNA_g4359.t1 and Bju-novel-miR173-A06_VARUNA_g238.t1/B04_VARUNA_g5283.t1/A05_VARUNA_g4026.t1 were predicted as key regulators of cell cycle and cell division, embryo development, triacylglycerol (TAG) synthesis, seed coat color, and seed size [10].

3.6.3. Group E3: miRNAs upregulated only in the late stage of EH2

A total of 144 unique target transcripts were mapped to miRNAs Bju-miR156.12 and Bju-miR390.1 (Supplementary Table 9) and comprised 144 miRNA-target modules. Among these, 121 miRNA-target transcript pairs showed an inverse correlation, and 16 pairs showed DE transcripts.

The TF enrichment showed enrichment of GRF TFs (Fig. 4c, Supplementary Table 12), and one of the target transcripts (A07_VARUNA_g4239.t1) of Bju-miR156.12 was identified as an important candidate as it is a putative regulator of cell cycle and cell division [10].

3.6.4. Group P1: miRNAs upregulated in PJK at both early and late stages

The target analysis showed 1079 unique target transcripts mapped to these miRNAs comprising 1082 miRNA-target modules, with 590 showing inverse correlation and 87 modules showing inverse differential expression pattern (Supplementary Table 9).

Pathway mapping revealed development, hormone metabolism, secondary metabolism, signaling, regulation of transcription, and DNA synthesis, etc. as enriched categories in this group (Fig. 4b, Supplementary Table 10). GO categories like cell cycle, cell division, embryo/post-embryonic development, seed development, etc. were also enriched in this dataset (Supplementary Table 11) and TFs like AP2, ARF, BBR-BPC, GRF, NAC, NF-YA, and SBP were overrepresented (Fig. 4c, Supplementary Table 12).

Some of the crucial targets like A09_VARUNA_g467.t1, A05_VARUNA_g2899.t1, A05_VARUNA_g1616.t1, and A10_VARUNA_g3003.t1, were mapped to embryo development, seed coat mucilage, cell cycle and cell division related genes [10]. Moreover, several miRNAs of this group, like miR156, miR164, miR167, miR169, miR172, and miR396 are important for regulating seed size and oil content during seed development [29,38,42,59,75,79,80].

3.6.5. Group P2: miRNAs upregulated in PJK at early stage only

There were 849 unique target transcripts for 26 miRNAs with a total of 1001 miRNA-target modules in this group, and 74 miRNA-target modules showed an inverse differential expression (Supplementary Table 9). Pathway mapping revealed cell division, development, lipid metabolism, synthesis, signaling, regulation of transcription, synthesis, processing, and transport (Fig. 4b, Supplementary Table 10), while GO enrichment delineated cytokinin hormone, post-embryonic development, and phase transition, as important categories (Supplementary Table 11). AP2, ARF, BBR-BPC, GRF, NAC, NF-YA, and SBP were overrepresented TFs in this group (Fig. 4c, Supplementary Table 12).

Some of the important targets, such as A05_VARUNA_g1616.t1, A06_VARUNA_g3049.t1, MSTRG.38655.2 B05_VARUNA_g6233.t1 and B04_VARUNA_g2725.t1 were putatively involved in regulating embryo development, cell cycle, and cell division, seed size and glucosinolate content [10].

3.6.6. Group P3: miRNAs upregulated in PJK at late stage only

The target analysis showed 404 unique target transcripts for 13 miRNAs, 425 miRNA-target modules with 22 pairs differentially expressed and inversely correlated (Supplementary Table 9). Pathway mapping revealed development, secondary metabolism, protein synthesis, sulphate transport, etc. as important categories (Fig. 4b, Supplementary Table 10). Moreover, GRAS, SBP, and VOZ TF families were overrepresented in this group (Fig. 4c, Supplementary Table 12). Some crucial modules like Bju-miR395.4-MSTRG.11621.2/A05_VARUNA_g2653.t1 were identified as regulators of the seed development [10].

We further shortlisted 36 differentially expressed miRNAs that were candidates for seed development roles and used all their target transcripts (2048) to identify the miRNA-target networks involved in regulating seed traits. Among these miRNA-target pairs, 40 unique targets were regulated by more than one miRNA family (Fig. S4). Several hub genes were identified in the network. For example, many transcripts of genes *PHABULOSA* (*PHB*), *PHAVOLUTA* (*PHV*), and *REVOLUTA* (*REV*) were common to miR165 and miR166. Another hub gene, *Nuclear Factor YA9* (*NF-YA9*, B01_VARUNA_g2960.t1), an important candidate for seed size, was targeted by both miR164 and miR169. The networks thus identified shall be helpful to investigate further the complexities in miRNA-mediated regulation of seed development in Brassica.

3.7. Discovery of candidates for regulation of seed traits

3.7.1. Seed size

Seed size is under complex regulatory control exerted through transcription factors, hormones, *HAIKU (IKU)* pathway, ubiquitin-proteasome pathway, various signaling pathways, and noncoding RNAs [10,14]. These players act in concert to regulate cell proliferation and cell enlargement in the seed coat, endosperm, and embryo [28]

(Fig. 5). We detected 15 known miRNAs in our data with characterized roles in the regulation of seed size/weight (Supplementary Table 13). Of these, 7 exhibited differential expression among both varieties. Further, a candidate gene search from the literature showed that 29 unique targets mapped to 13 seed-size candidate genes (Fig. 5, Supplementary Table 13). These were targeted by 22 miRNAs, of which seven were differentially expressed. Some of these; for example, Bju-miR172.2 and Bju-novel-miR189 that showed upregulation in PJK and targeted negative regulators, *Apetala 2 (AP2)* and *Arabidopsis Histidine Kinase 2 (AHK2),* are excellent candidates.

3.7.2. Seed coat color

Seed coat color is regulated by complex gene regulatory networks of flavonoid metabolic pathways that determine the content of flavanols, anthocyanins, and pro-anthocyanidins [55]. These networks are well characterized in *Arabidopsis* and this information helps pick candidates for seed coat color in Brassicas as well [7]. The early biosynthetic genes like *Chalcone Synthase (CHS), Chalcone Isomerase (CHI), Flavanone 3-Hydroxylase (F3H),* and *Flavanone 3′-Hydroxylase (F3′H)* are involved in the production of precursors that are finally converted into proanthocyanidins through a series of steps (Fig. 6) [7]. Based on target mapping to genes that function in the flavonoid biosynthetic pathway, we found 19 unique target transcripts in our data corresponding to 12 flavonoid biosynthetic genes (Supplementary Table 14). These were targeted by 15 unique miRNAs, comprising 25 miRNA-target modules. Of these, seven were differentially expressed (Fig. 6, Supplementary Table 14). Notably, miR156, which putatively targeted flavonoid pathway gene *TRANSPARENT TESTA 1* (*TT1*), was upregulated in EH2, while miR395, which targeted gene *FLAVONOL SYNTHASE 1* (*FLS1*), was upregulated in PJK [7].

3.7.3. Oil content

We detected various miRNAs involved in oil biosynthesis (Supplementary Table 15). For instance, miR156 has been shown to exhibit peak expression during the middle phase of seed development in *B. napus*. It has been predicted as a critical candidate for regulating fatty acid biosynthesis [64]. The expression pattern of miR159 strongly correlates with oil content in various varieties of *B. napus*. It has been associated with fatty acid biosynthesis in *B. napus* [64]. Zhao et al., [78] also deduced miR156 and miR167 as possible regulators of oil biosynthesis in *B. napus*.

Next, we carried out triacylglycerol (TAG) biosynthesis pathway reconstruction using available literature. TAG biosynthesis primarily occurs through the highly conserved Kennedy pathway, which converts glycerol-3-phosphate into triacylglycerols [15]. We collated the knowledge of metabolites and enzymes involved in fatty acid synthesis in the plastids and accumulation of TAG in the endoplasmic reticulum (Fig. 7) as reported previously [41,64]. We observed that 18 of the transcripts expressed in *B. juncea* seeds are putatively targeted by seven miRNAs detected in our data (Fig. 7, Supplementary Table 15). We determined that 18 miRNA-target modules are candidates for regulating TAG biosynthesis in *B. juncea*. Of these, two miRNAs, Bju-miR169.9 and Bju-miR858.2, were differentially expressed. miR169, which targeted gene beta-ketoacyl-ACP- synthase II (KASII), was upregulated in PJK, while miR858, which targeted phospholipase D (PLD), was upregulated in EH2. *KAS II* carries out elongation of 16:0 acyl carrier protein (ACP) to 18:0 ACP, and mutation in the KASII gene modulates the oil composition of *Arabidopsis* seeds [48]. *Phospholipase D* (*PLD*) carries out hydrolysis of phosphatidylcholine (PC) into phosphatidic acid (PA), which is next converted into diacylglycerol (DAG), and is therefore an important determinant of TAG flux [69].

3.8. Co-localization of seed weight and oil content miRNA candidates with QTLs

We used previously available information on seed weight (*Tsw*) and oil content (*Oil*) QTL from a bi-parental population EPJ derived from EH2 and PJK varieties [12,50] and identified differentially expressed miRNAs or their targets co-localizing with the genomic intervals harboring these QTLs (Fig. 8).

For seed weight, we identified 29 unique miRNAs colocalized with 11 QTLs. Among these miRNAs, 16 were differentially expressed. MiRNAs like Bju-miR156.12, Bju-miR156.13, Bju-miR319.3, Bju-miR390.1, Bju-novel-miR154, and Bju-novel-miR174 were upregulated in EH2, while Bju-miR156.16, Bju-miR156.18, Bju-miR157.2, Bju-miR395.2, Bju-miR395.3, Bju-miR395.4, Bju-miR399.4, Bju-novel-miR87, and Bju-novel-miR163 were upregulated in PJK (Supplementary Table 16). Further, QTL information showed that in the study by Dhaka et al., [13], all positive alleles for seed weight were contributed by the parent PJK. Therefore, the identification of miRNAs that act as positive alleles for seed weight regulation showed that miRNAs Bju-miR167.3 and Bju-miR319.3, which fall within *EPJ-Tsw-B1–1* and *EPJ-Tsw-A3–1*, respectively, are good candidates. Bju-miR167.3 is upregulated, while Bju-miR319.3 is downregulated in PJK. MiR167 positively regulates the

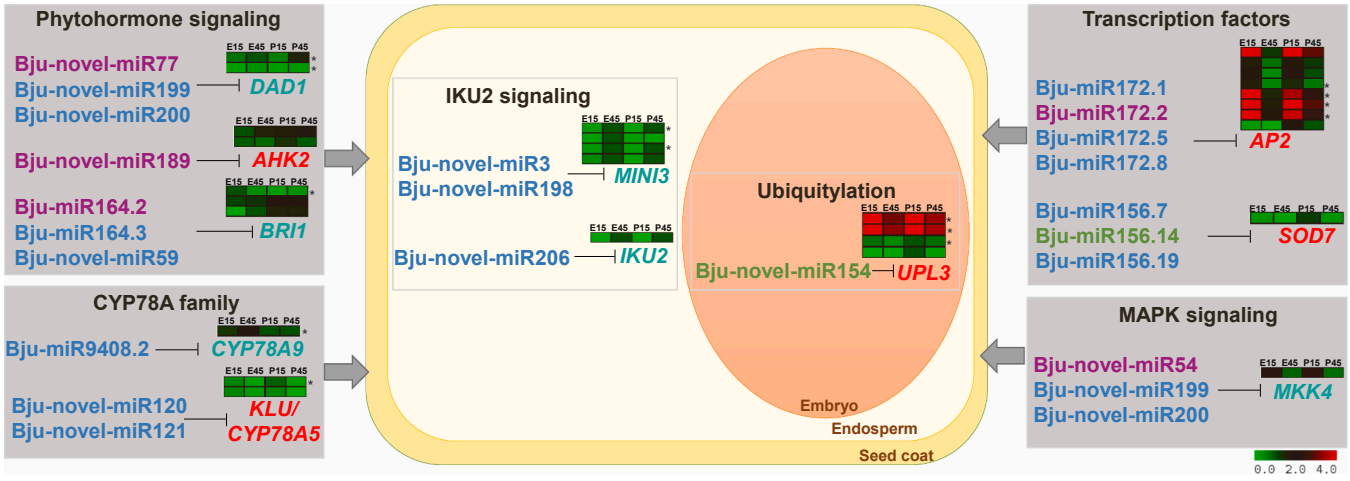

Fig. 5. MiRNA candidates for seed size determination in *B. juncea*. Genes are orthologous to target transcripts with experimentally proven roles in regulating cell proliferation or cell expansion in the seed coat, endosperm, or embryo. Positive regulators are shown in teal, and negative regulators are shown in red. MiRNAs respectively targeting such characterized genes are also shown (blue– not differentially expressed, green – upregulated in EH2, purple – upregulated in PJK, * represents inversely correlated targets). Heatmaps showing expression levels of the transcripts orthologous to target genes represented as $\log_2$ (TPM+1) are also shown.

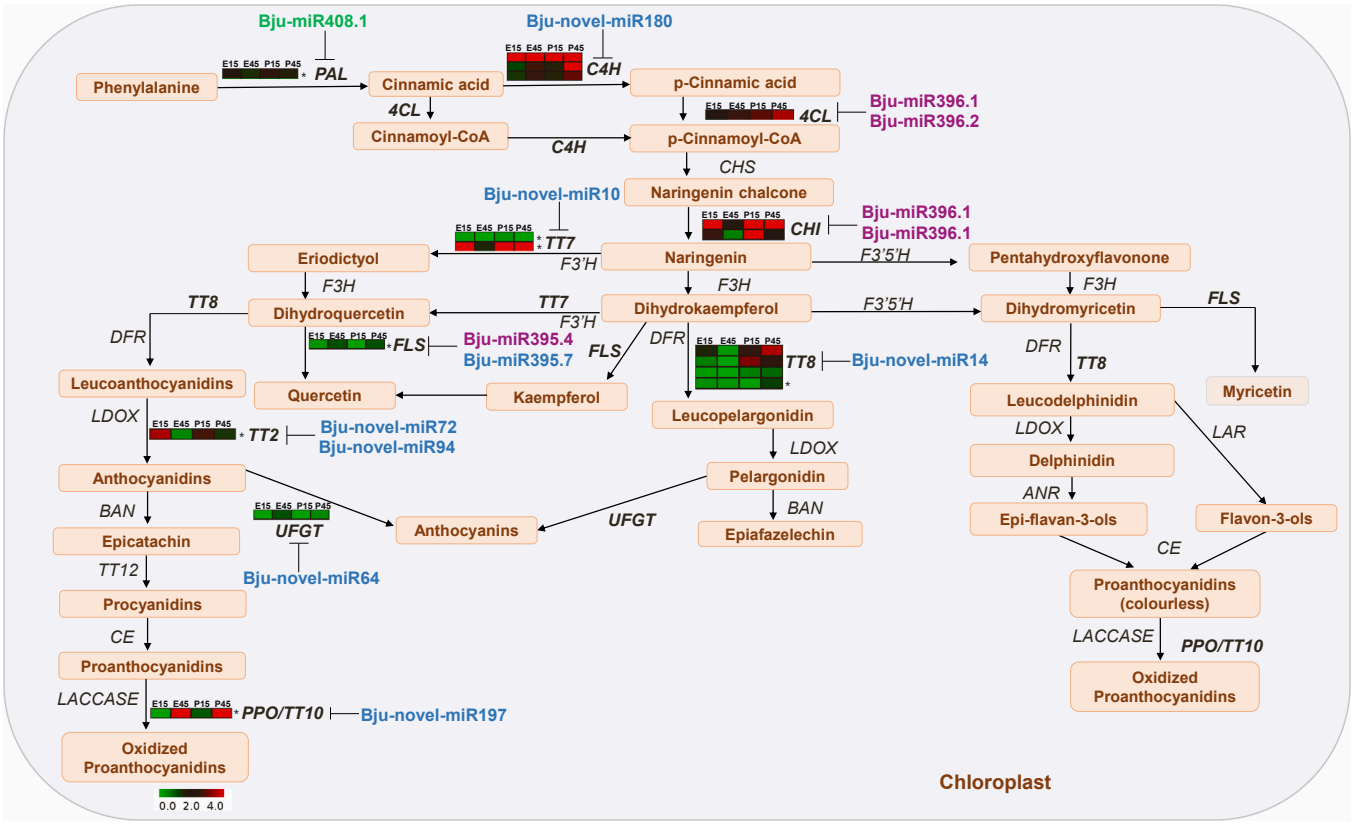

Fig. 6. MiRNA candidates involved in seed coat color determination in *B. juncea*. The flavonoid biosynthetic pathway information was retrieved from the literature [7,17,41,47,55,77]. The genes orthologous to predicted targets are shown in bold, and respective miRNAs (blue– not differentially expressed, green – upregulated in EH2, purple – upregulated in PJK, * represents inversely correlated targets) detected in *B. juncea* are also shown. Heatmaps showing expression levels of the transcripts orthologous to target genes represented as $\log_2$ (TPM+1) are also shown. Genes abbreviations: *PAL, phenylalanine ammonia lyase; C4H, cinnamate 4 hydroxylase; 4CL, 4-coumarate coenzyme A ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3- hydroxylase; F3′H, flavanone 3′-hydroxylase; F3′5′H, flavonoid 3′,5′-hydroxylase; FLS, flavonol synthase; DFR, dihydroflavonol-4-reductase; LDOX, leucoanthocyanidin dioxygenase; ANR, anthocyanidin reductase; BAN, BANYULS; LAR, leucoanthocyanidin reductase; TTG1, transparent testa glabra1; CE, condensing enzyme; UFGT, UDP flavonoid glycosyl transferase; TT2/7/8/10, transparent testa 2/7/8/10; PPO, polyphenol oxidase.*

seed size [38], while miR319 negatively regulates the cell proliferation [52]. Another interesting observation was mapping the same miRNAs to different seed-weight QTL regions; for example, Bju-miR390.1 was present within the QTL regions on both A03 and B03 chromosomes.

We identified 15 unique miRNAs for oil content, including nine differentially expressed miRNAs colocalizing with seven QTLs. Bju-miR169.9 was a likely causal gene candidate as it targets TAG biosynthesis gene *KAS II*. It was upregulated in PJK, and the positive alleles for QTLs *Oil-B7–2-EPJ* and *Oil-B7–2-EPJ* are also contributed by PJK. Two different members of miR319 coincided with oil QTLs on chromosomes A01 and A03, while miR169 family members co-located with oil QTLs on B01 and B07 (Supplementary Table 16).

3.9. Validation of differentially expressed miRNAs and their targets by qPCR

To validate the findings of our small RNA sequencing experiment, we randomly selected six candidate miRNAs and their respective predicted target transcripts (miR172-A03_VARUNA_g6127.t1, miR319-A01_VARUNA_g993.t1, miR390-A01_VARUNA_3145.t1, miR394-A08_VARUNA_g3145.t1, miR408-B08_VARUNA_g3369.t1, and miR858-B07_VARUNA_g5342.t1) for qRT-PCR analysis. All these six miRNAs exhibited similar patterns as observed in the small RNA sequencing experiment. Of these, miR172 and miR394 exhibited upregulation in PJK, while miR319, miR390, miR408, and miR858 showed upregulation in EH2. Further, target transcripts of three of these, miR172, miR394, and miR408, showed inverse expression patterns, confirming the robustness of our sequencing data.

4. Discussion

4.1. Conserved characteristics of B. juncea seed miRNAome

Seed miRNAomes in oilseed Brassicas have been previously reported in *B. napus* [19,24,49,58,64,66,67,78]. In the case of *B. juncea*, we recently examined the seed transcriptome dynamics. We identified early (15D) and late (45D) development stages of the seed as transcriptionally distinct [10]. We also determined that the contrast in the expression of cell division-related genes may be a significant regulator of seed size in the varieties PJK and EH2. In this study, we have integrated our previous findings with the miRNA profiles generated using the early and late stages of seeds from EH2 and PJK to identify key miRNAs regulating seed size and other seed traits in *B. juncea*. The *B. juncea* seed miRNAome thus identified comprised 326 miRNAs, including 127 known and 199 novel miRNAs, respectively (Fig. 1a, Supplementary Table 3). Most of the mature miRNAs (68%) in our findings belonged to the 21-nt class, confirming the typical length distribution of miRNAs in most plant species (Fig. 1d & e) [2]. Additionally, the prevalence of uracil at 5′ positions of miRNAs (57%) was consistent with the involvement of DCL1 in plant miRNA biogenesis [16]. The structural features of miRNA precursors with MFE ranging from − 36 to − 184 Kcal/mol further confirmed the robustness of our miRNAome identification (Supplementary Table 4). The genomic localization confirmed that most of the miRNAs (76%) were intergenic, a finding consistent with previous

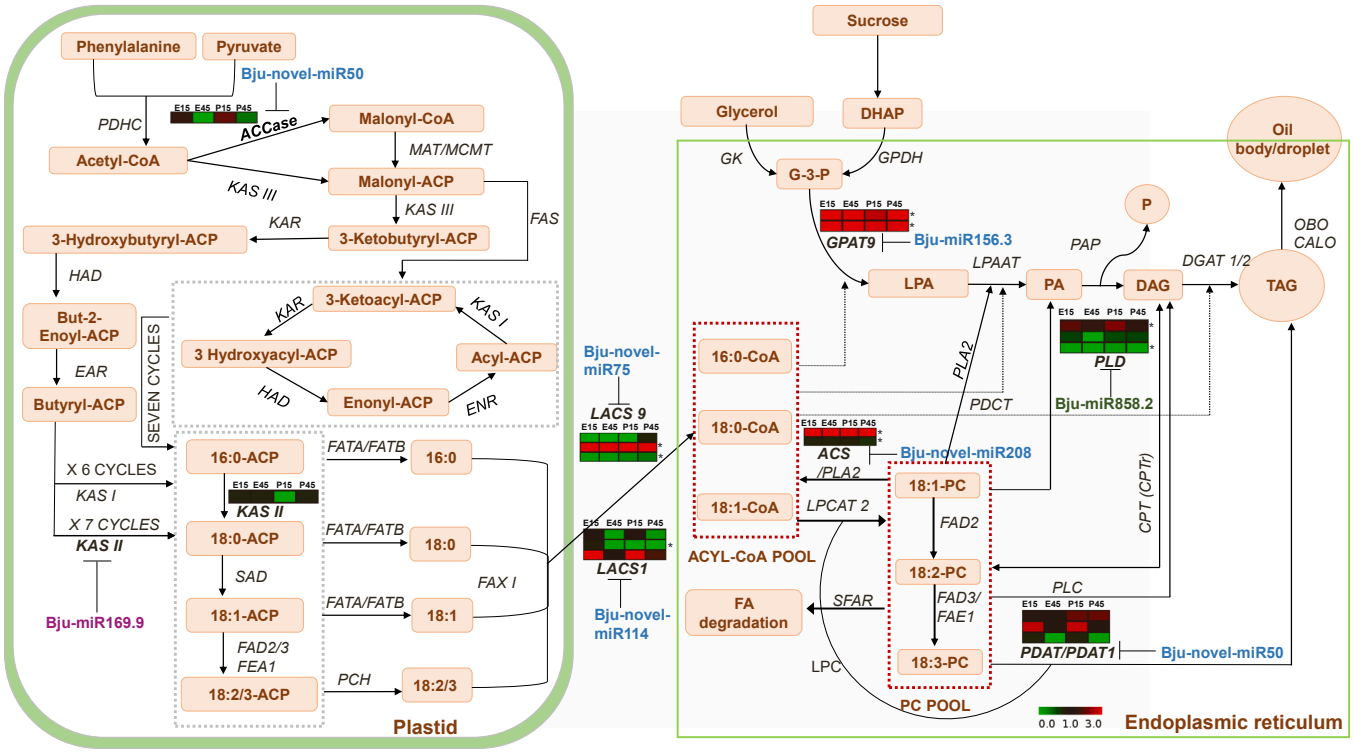

Fig. 7. MiRNA candidates involved in oil content determination in *B. juncea*. The oil biosynthetic pathway was reconstructed using information from the literature [41,51,64,65,76]. Formation of the fatty acid from the carbon source occurs in the plastid. TAG (Triacylglycerol) assembly/synthesis from the phosphatidylcholine-derived pathway and Kennedy pathway occur in the endoplasmic reticulum. Orthologs of targets are shown in bold, and corresponding miRNAs (blue– not differentially expressed, green – upregulated in EH2, purple – upregulated in PJK, * represents inversely correlated targets) are also shown. Heatmaps showing expression levels of the transcripts orthologous to target genes represented as $\log_2$ (TPM+1) are also shown. Protein and FAs abbreviations: *ACCase, acetyl-CoA carboxylase*; PDHC, pyruvate dehydrogenase complex; ACP, acyl carrier protein; MCMT, malonyl-CoA *malonyltransferase*; *MAT, malonyl-CoA-ACP transacylase*; *PLD, phospholipase D*; *KAS III, 3-ketoacyl-ACP synthase III*; *KAS II, 3-ketoacyl-ACP synthase II; HAD, hydroxyacyl-ACP dehyroxyacyl-ACP Dehydrase; KAR, ketoacyl-ACP reductase*; *ENR, enoly- ACP reductase*; FATA/B, fatty acyl-ACP thioesterase A/B; *GPAT9, glycerol-3-phosphate acyltransferase 9*; *SAD, stearoly-acyl carrier protein desaturase*; FAX1, *plastid fatty acid export 1*; *LACS9, long-chain acyl-CoA synthetase 9*; *DGAT, diacyltransferas*; *LPCAT, lysophosphatidylcholine acyltransferase*; *GPDH, glycerol-3- phosphate dehydrogenase*; *LPAAT, lysophosphatidic acid acyltransferase*; *PAP, phosphatidic acid phosphatase*; *FAD2/3, fatty acid desaturase 2/3*; FAS, fatty acid synthase; *PDCT, phosphatidylcholine: diacylglycerol cholinephosphotransferase*; *PDAT, phospholipid: diacylglycerol acyltransferase*; PA, Phosphatidic acid; *CPT CDP, choline: diacylglycerol cholinephosphotransferase*; *PLC, phospholipase C*; DHAP, dihydroxy acetone phosphate; *CALO, caleosin*; *OBO, oil body oleosin,* steroleosin (oil proteins); *SFAR, seed fatty acid reducer*; G-3-P, glycerol-3- phosphate; LPA, lyso-phosphatidic acid; DAG, diacylglycerol; TAG, triacylglycerol; PC, phosphatidylcholine; *PLA2, phosphalipase A2* (beta); *FAE 1, fatty acid elongase*; *EAR, enoyl-ACP reductase*; *ACS, Acyl CoA Synthetase;* LPC, lysophophatidylcholine.

reports (Fig. S3a) [43].

4.2. *Highly expressed miRNAs are crucial for seed development*

We observed the miR156 family has the highest number of family members in our data, and target analysis using our transcriptome data confirmed a total of 397 miR156-SPL modules (Supplementary Table 3, 8). MiR156 was also the highest expressed miRNA in our data (Supplementary Table 5). Interestingly, miR156 was also reported to be the most prominent family with 115 members in the seed miRNAome of *B. napus* [19]. They also confirmed miR156 as the topmost expressed miRNA, contributing to more than 48% of the total miRNA reads. Other Brassica seed miRNAome studies also reported miR156 as one of the most critical candidates for regulating the seed development [24,58,64, 66,67]. Abundant expression of miR156 throughout the developing embryos has been previously shown in *Arabidopsis*. MiR156-SPL module regulates seed maturation [19]. Hence, miR156 is likely a central player in *B. juncea* seed development.

Similarly, miR158 was among the most abundant miRNA families in *B. napus* seed [19]. MiR159 was preferentially expressed in *B. napus* embryos and is predicted to be involved in ABA-mediated seed maturation [19]. There is also a high expression of miR160, miR166, and miR167, etc. in the *B. napus* seed development [19]. These observations imply that the seed miRNAomes appear relatively conserved across *Brassica* species. Some of the other highest expressed known miRNAs in our data, like miR160, miR165, miR166, miR167, and miR319, have proven roles in regulating the embryogenesis [31,37,52,70]. Interestingly, we found only two known miRNAs, miR400 and miR5654, among the top expressed miRNAs, which have no proven roles in seed development yet. These are new candidates to investigate their roles in seed development. MiR400 is expressed highly in both stages in both varieties. It has 45 putative targets, mostly Pentatricopeptide repeat (PPR) superfamily proteins, known to participate in the abiotic stress response [34]. MiR5654 had two members and 40 putative targets, most of which were Tetratricopeptide repeat (TPR)-like superfamily proteins. Both miR400 and miR5654 were also reported to express during silique development in *B. napus* [6].

Interestingly, most of these miRNAs exhibited similar expression levels in all four categories (E15, E45, P15, P45), except miR156.8 and miR319.2, which showed low expression in P15 compared to the other stages. While most of the top expressed miRNAs showed comparable expression levels in both EH2 and PJK (Fig. 2b), Bju-miR156.8, the miRNA with the highest average expression in our data, showed high differential expression among EH2 and PJK. It showed $>$ 8-fold upregulation in EH2 at 15D compared to PJK. Similarly, Bju-miR319.2 also showed $>$ 5-fold upregulation in EH2. Thus, Bju-miR156.8 and Bju-miR319.2 are important candidates for investigating inter-varietal differences in *B. juncea* seed development.

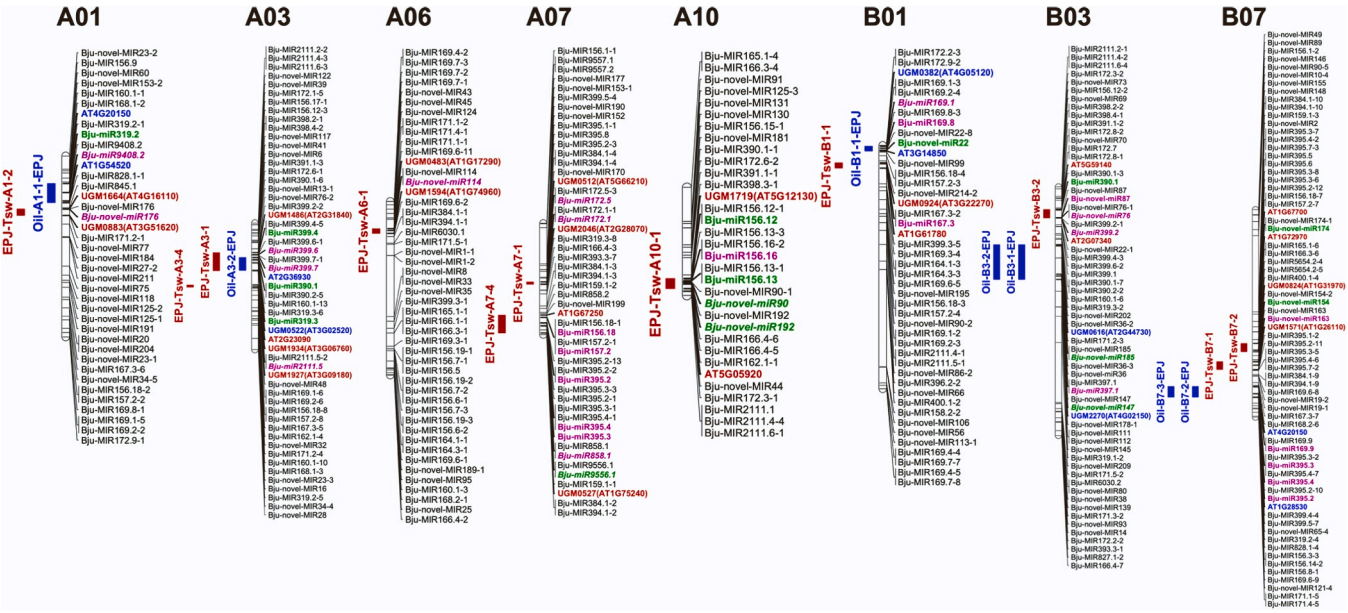

Fig. 8. Co-localization of miRNAs and seed weight and oil content QTLs in *B. juncea*. Color coding for miRNAs: green- upregulated or exclusively expressed in EH2, purple-upregulated or exclusively expressed in PJK. Color coding for QTLs: red- seed weight QTLs and their flanking markers, blue- oil content QTLs along with their flanking markers.

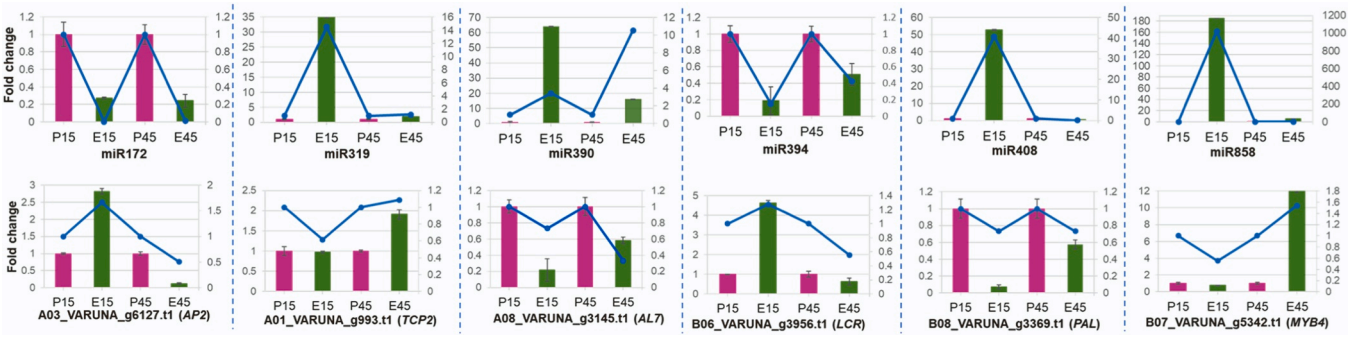

Fig. 9. Expression validation of differentially expressed miRNAs and their corresponding target genes by qRT-PCR. The top panel shows the qRT-PCR results for six miRNAs with green bars for EH2 and purple bars for PJK. The blue line depicts the fold changes obtained from the RNA-sequencing data. The bottom panel shows the qRT-PCR results for the corresponding target genes.

4.3. Comparative functional analysis of miRNA dynamics reveals candidates for improving seed traits

The comparison of variety-wise miRNA expression showed that 211 miRNAs were expressed in a variety-specific manner, with 95 and 116 miRNAs expressing specifically in one of the varieties only. Of these, 36 miRNAs were expressed in both stages in EH2 but not in PJK, while 20 miRNAs were expressed in both stages of PJK but not in EH2. These miRNAs are good candidates for investigating variety-specific differences in seed traits. Further analysis of this set of miRNAs showed that 12 and 32 were the EH2- and PJK-specific known miRNAs (Supplementary Table 6).

Among the 103 differentially expressed miRNAs, 43 and 24 were upregulated in EH2 at the early and late stages, respectively, while 45 and 32 were upregulated in PJK at the early and late stages (Supplementary Table 7). A more significant number of miRNAs were differentially expressed at the early stage compared to the late stage. Further, dividing miRNAs into different groups based on their stage-wise and variety-wise differential expression patterns enabled in-depth analyses of our data. We observed that different miRNA family members of miR156, miR158, and miR166 were upregulated in both varieties, suggesting that these miRNAs may perform the same function despite allelic differences. On the other hand, several miRNAs showed variety-specific upregulation and were, therefore, suggestive of their roles in regulating seed traits in a variety-specific manner. Integrating these leads with the transcriptome analysis using target identification, inverse expression correlation, and candidate gene information further expanded our search for valuable candidates from both EH2 and PJK.

4.3.1. Beneficial miRNA candidates from small and yellow-seeded EH2

Expression analysis identified miR9556 as an EH2-specific family. Its function is unknown, but a previous study in *B. napus* has suggested its involvement in the phenylpropanoid metabolism pathway [40], indicating that it may be important for regulating seed coat color. Further, groupwise demarcation of differentially expressed miRNA families highlighted miR161, miR165, miR319, miR390, miR403, miR408, miR827, miR858, and miR6030 as specifically upregulated in EH2. Some of these miRNAs have previously characterized roles in regulating seed traits. For instance, miR319 is known to repress TCP4, a repressor of cell proliferation [52]. In wheat, the miR319 GAMYB module regulates cell proliferation, grain weight, and yield [21]. Hence, its upregulation in EH2 implies low cell proliferation during seed development in the small-seeded variety. In chickpea, miR319 levels were higher in the small-seeded variety [23].

Further, several members of the miR319 family exhibited colocalization with a seed weight QTL on the A03 chromosome and oil QTLs on the A01 and A03 chromosomes (Fig. 8, Supplementary Table 16). miR827 is upregulated in the short silique line of *B. napus* [6], and due to its role in phosphate starvation response, the authors suggested that it may be upregulated to enable enhanced phosphate uptake in small siliques. Thus, it may serve the same function in *B. juncea* as well.

In the late stage, miR390 was upregulated exclusively in EH2. MiR390 has been proposed to regulate early embryo development through auxin-mediated signaling in *B. napus* [78]. Besides, the Bju-miR390.1-A02_VARUNA_g716.t1 module in group E1 exhibited an inverse differential expression pattern. This target transcript is *TRANSPARENT TESTA 19 (TT19)* ortholog, a positive regulator of flavonoid accumulation in seed coat [30]. *BjuTT19* which encodes for glutathione transferase is experimentally shown to be highly downregulated in yellow-seed variety and highly expressed in brown seeded variety of *B. juncea* [55]. Therefore, upregulation of miRNA targeting it in EH2 suggests involvement in regulating seed coat color. MiR390 also exhibited colocalization with two seed-weight QTLs on different chromosomes (Fig. 8, Supplementary Table 16).

Another critical candidate from EH2 was miR858, a negative regulator of the flavonoid synthesis pathway, as shown in *Arabidopsis* [54]. Its overexpression leads to repression of MYB TFs involved in promoting flavonoid biosynthesis. Moreover, Bju858.2 also targets an oil biosynthetic gene *PLD* (B04_VARUNA_g4318.t1, MSTRG.36636.2, and MSTRG.36636.3) and hence may be a target for modulating seed coat color as well as oil content.

We also noted that not all characterized miRNAs followed the expected patterns. For example, miR403 has been shown to target *Argonaute 2 (AGO2)*. It shows higher expression in the high seed weight variety and is positively correlated with seed yield *in B. napus* [74], while we observed its high expression in the low seed weight variety. Similarly, miR408 positively impacts the seed size [45], while we observed an opposite trend (Supplementary Table 9). These patterns may reflect species-specific variations. However, miR408 was also detected as putatively targeting (B08_VARUNA_g3369.t1), an ortholog of *Phenylalanine ammonia-lyase (PAL)* (Supplementary Table 14). *PAL* is crucial in flavonoid metabolism pathway and it is downregulated in yellow seeded line and upregulated in black seeded line in *B. napus* [39].

4.3.2. *Beneficial miRNA candidates from bold and brown-seeded PJK*

MiRNA expression showed that miR397, miR828, miR860, and miR2111 were PJK-specific families and nine families, miR157, mir164, miR167, miR169, miR172, miR393, miR394, miR395 and miR399 showed upregulation exclusively in PJK. Deeper insights into the functions of these families revealed that miR828 is involved in anthocyanin accumulation. In grape, miR828 has been experimentally shown to repress the MYBs that act as anthocyanin accumulation repressors [61]. Therefore, our finding that miR828 is specific to the brown seed coat color variety PJK and is absent in the yellow seeded variety EH2 is highly promising. MiR2111 had six members, but none expressed in EH2. A higher prevalence of miR2111 was observed in the high oil-content variety of *B. napus* in two separate studies [64,66], indicating that it might be an important candidate for regulating oil content.

Several of the PJK-upregulated miRNAs were positive regulators of seed size. Overexpression of miR167 in *Camelina* enhances its seed size and oil content, [38], and is therefore a good candidate for seed size enhancement. Similarly, miR169 overexpression enhances seed size and weight in maize [73]. In the case of miR172, overexpression reduced seed weight in rice but increased seed size in jatropha seed [59,80]. MiR393 is also likely a positive regulator of seed size as overexpression of its target gene reduces seed size in cucumber [68]. Still, the opposite pattern is observed in rice [4]. MiR394 regulates seed size positively as its overexpression enhances seed size [56].

Besides these candidates, Bju-miR172.2 and Bju-novel-miR189 were also crucial as they were upregulated in PJK, and they putatively targeted negative regulators of seed size, *AP2*, and *AHK2*, respectively. *AP2* gene represses cell enlargement in seed coat and cell proliferation in the endosperm [44]. *AHK2* codes for cytokinin receptors, and their mutation leads to an enhanced seed size [3]. MiR169 also targeted B05_VARUNA_g1925.t1, an ortholog of KAS II, a catalytic enzyme in the fatty acid pathway.

5. Conclusions and future perspectives

We utilized genetically diverse and contrasting varieties EH2 and PJK to characterize the seed miRNAome from two distinct stages of seed development in *B. juncea*. Using a detailed investigation of the miRNA expression patterns, target analyses, and previously generated transcriptome and QTL data, we identified robust candidates for improving economically important seed traits in *B. juncea*. Identification of variety-specific miRNAs is highly beneficial as ubiquitous miRNAs expressed at high levels likely regulate multiple functions in a conserved manner and cannot be targets of engineering the seed traits. On the other hand, variety-specific targets likely contribute to unique traits, and their modulation is less likely to negatively impact overall growth and yield. Hence, we propose miR167, miR169, and miR319 as the best candidates for improvement of seed size and oil content, miR172 and miR394 for improvement of seed size, miR390, miR408, miR828, and miR858 for the improvement of seed coat color. Besides these conserved miRNAs, we detected several other known and novel miRNAs regulating these traits. These miRNAs and target genes are suitable for marker development and future experimental validations toward enhancing the seed size, oil yield, and quality in *B. juncea*.

Supplementary figure legends

Fig. S1**: Size distribution of small RNAs.**

Fig. S2**: Genomic localization of miRNA precursors.** Blue boxes show different family members of the same miRNA family on the same chromosome.

Fig. S3**: Characteristics of miRNA precursors.** (a) Distribution of miRNAs shown in different genomic locations. (b) Distribution of minimum free energy per nucleotide length (MFE/nt) of miRNA precursors.

Fig. S4**: Network analysis of differentially expressed miRNAs involved in seed development and their target transcripts.** Diamonds represent miRNAs, circles represent target transcripts, and rectangles represent common target transcripts shared by multiple miRNA families.

CRediT authorship contribution statement

Conceptualization: ND; Formal analysis: RJ, ND, PY, MD, SK, SS; Funding acquisition: ND; Methodology: RJ, IV; Supervision: ND, MKS, RBS, RS; Visualization: RJ, ND, PY, MD, SK, SS; Writing – original draft: RJ, ND; Writing – review & editing: ND, RJ, MKS, RBS, RS. All authors read and approved the final manuscript.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Namrata Dhaka reports financial support was provided by India Ministry of Science & Technology Department of Science and Technology.

Data Availability

The small RNA sequencing datasets generated in this study have been submitted to the NCBI Sequence Read Archive (SRA) under the Bioproject:PRJNA1004201.

Acknowledgements

ND thanks research grants from the Department of Science and Technology Government (DST) (Research Grant IFA17-LSP90 (DST INSPIRE)) and Science and Engineering Research Board (SERB), (CRG/2019/001695), Government of India. RJ acknowledges financial assistance through an ICMR-SRF fellowship. PY acknowledges the University Non-NET fellowship from the Central University of Haryana, Mahendergarh. IV thanks CSIR for CSIR-RA fellowship. The authors also thankful for the facilities provided by Jawaharlal Nehru University, New Delhi and Central University of Haryana, Mahendergarh. The seeds for this study were kindly provided by Prof. Akshay Kumar Pradhan, Department of Genetics, University of Delhi, South Campus, New Delhi.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cpb.2023.100306.

References

[1] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[2] M.J. Axtell, B.C. Meyers, Revisiting criteria for plant microRNA annotation in the era of big data, Plant Cell 30 (2018) 272–284.
[3] I. Bartrina, H. Jensen, O. Novák, M. Strnad, T. Werner, T. Schmülling, Gain-of-function mutants of the cytokinin receptors AHK2 and AHK3 regulate plant organ size, flowering time and plant longevity, Plant Physiol. 173 (2017) 1783–1797.
[4] H. Bian, Y. Xie, F. Guo, N. Han, S. Ma, Z. Zeng, J. Wang, Y. Yang, M. Zhu, Distinctive expression patterns and roles of the miRNA393/TIR1 homolog module in regulating flag leaf inclination and primary and crown root growth in rice (*Oryza sativa*), N. Phytol. 196 (2012) 149–161.
[5] R. Chandna, R. Augustine, N.C. Bisht, Evaluation of candidate reference genes for gene expression normalization in *Brassica juncea* using real time quantitative RT-PCR, PloS One 7 (2012), e36918.
[6] Li Chen, Lei Chen, X. Zhang, T. Liu, S. Niu, J. Wen, B. Yi, C. Ma, J. Tu, T. Fu, Identification of miRNAs that regulate silique development in *Brassica napus*, Plant Sci. 269 (2018) 106–117.
[7] Y.-Y. Chen, H.-Q. Lu, K.-X. Jiang, Y.-R. Wang, Y.-P. Wang, J.-J. Jiang, The flavonoid biosynthesis and regulation in *Brassica napus*: a review, Int. J. Mol. Sci. 24 (2022) 357.
[8] X. Dai, Z. Zhuang, P.X. Zhao, psRNATarget: a plant small RNA target analysis server (2017 release), Nucleic Acids Res. 46 (2018) W49–W54.
[9] Á. Dalmadi, P. Gyula, J. Bálint, G. Szittya, Z. Havelda, AGO-unbound cytosolic pool of mature miRNAs in plant cells reveals a novel regulatory step at AGO1 loading, Nucleic Acids Res. 47 (2019) 9803–9817.
[10] N. Dhaka, R. Jain, A. Yadav, P. Yadav, N. Kumar, M.K. Sharma, R. Sharma, Transcriptome analysis reveals cell cycle-related transcripts as key determinants of varietal differences in seed size of *Brassica juncea*, Sci. Rep. 12 (2022), 11713.
[11] N. Dhaka, K. Krishnan, M. Kandpal, I. Vashisht, M. Pal, M.K. Sharma, R. Sharma, Transcriptional trajectories of anther development provide candidates for engineering male fertility in sorghum, Sci. Rep. 10 (2020), 897.
[12] N. Dhaka, A. Mukhopadhyay, K. Paritosh, V. Gupta, D. Pental, A.K. Pradhan, Identification of genic SSRs and construction of a SSR-based linkage map in *Brassica juncea*, Euphytica 213 (1) (2017) 13.
[13] N. Dhaka, K. Rout, S.K. Yadava, Y.S. Sodhi, V. Gupta, D. Pental, A.K. Pradhan, Genetic dissection of seed weight by QTL analysis and detection of allelic variation in Indian and east European gene pool lines of *Brassica juncea*, Theor. Appl. Genet. 130 (2017) 293–307.
[14] N. Dhaka, R. Sharma, MicroRNA-mediated regulation of agronomically important seed traits: a treasure trove with shades of grey!, Crit. Rev. Biotechnol. 41 (2021) 594–608.
[15] F. dos Santos Maraschin, F.R. Kulcheski, A.L.A. Segatto, T.S. Trenz, O. Barrientos-Diaz, M. Margis-Pinheiro, R. Margis, A.C. Turchetto-Zolet, Enzymes of glycerol-3-phosphate pathway in triacylglycerol synthesis in plants: Function, biotechnological application and evolution, Prog. Lipid Res. 73 (2019) 46–64.
[16] A.L. Eamens, N.A. Smith, S.J. Curtin, M.-B. Wang, P.M. Waterhouse, The *Arabidopsis thaliana* double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes, Rna 15 (2009) 2219–2235.
[17] M. Hong, K. Hu, T. Tian, X. Li, L. Chen, Y. Zhang, B. Yi, J. Wen, C. Ma, J. Shen, Transcriptomic analysis of seed coats in yellow-seeded *Brassica napus* reveals novel genes that influence proanthocyanidin biosynthesis, Front. Plant Sci. 8 (2017) 1674.
[18] E. Howe, K. Holton, S. Nair, D. Schlauch, R. Sinha, J. Quackenbush, Mev: multiexperiment viewer, Biomed. Inform. Cancer Res. (2010) 267–277.
[19] D. Huang, C. Koh, J.A. Feurtado, E.W. Tsang, A.J. Cutler, MicroRNAs and their putative targets in *Brassica napus* seed maturation, BMC Genom. 14 (2013) 1–25.
[20] R. Jat, V. Singh, P. Sharma, P. Rai, Oilseed brassica in India: Demand, supply, policy perspective and future potential, OCL 26 (2019) 8.
[21] C. Jian, P. Hao, C. Hao, S. Liu, H. Mao, Q. Song, Y. Zhou, S. Yin, J. Hou, W. Zhang, The miR319/TaGAMYB3 module regulates plant architecture and improves grain yield in common wheat (*Triticum aestivum*), N. Phytol. 235 (2022) 1515–1530.
[22] I. Kalvari, E.P. Nawrocki, N. Ontiveros-Palacios, J. Argasinska, K. Lamkiewicz, M. Marz, S. Griffiths-Jones, C. Toffano-Nioche, D. Gautheret, Z. Weinberg, Rfam 14: expanded coverage of metagenomic, viral and microRNA families, Nucleic Acids Res. 49 (2021) D192–D200.
[23] N. Khemka, M. Singh Rajkumar, R. Garg, M. Jain, Genome-wide profiling of miRNAs during seed development reveals their functional relevance in seed size/weight determination in chickpea, Plant Direct 5 (2021), e00299.
[24] A.P. Koerbes, R.D. Machado, F. Guzman, M.P. Almerao, L.F.V. de Oliveira, G. Loss-Morais, A.C. Turchetto-Zolet, A. Cagliari, F. dos Santos Maraschin, M. Margis-Pinheiro, Identifying conserved and novel microRNAs in developing seeds of *Brassica napus* using deep sequencing, PloS One 7 (2012), e50663.
[25] A. Kozomara, M. Birgaoanu, S. Griffiths-Jones, miRBase: from microRNA sequences to function, Nucleic Acids Res. 47 (2019) D155–D162.
[26] Z. Kuang, Y. Wang, L. Li, X. Yang, miRDeep-P2: accurate and fast analysis of the microRNA transcriptome in plants, Bioinformatics 35 (2019) 2521–2522.
[27] B. Langmead, Aligning short sequencing reads with Bowtie, Curr. Protoc. Bioinforma. 32 (2010) 11–17.
[28] N. Li, R. Xu, Y. Li, Molecular networks of seed size control in plants, Annu. Rev. Plant Biol. 70 (2019) 435–463.
[29] S. Li, F. Gao, K. Xie, X. Zeng, Y. Cao, J. Zeng, Z. He, Y. Ren, W. Li, Q. Deng, The OsmiR396c-OsGRF4–OsGIF1 regulatory module determines grain size and yield in rice, Plant Biotechnol. J. 14 (2016) 2134–2146.
[30] X. Li, P. Gao, D. Cui, L. Wu, I. Parkin, R. Saberianfar, R. Menassa, H. Pan, N. Westcott, M.Y. Gruber, The Arabidopsis tt19–4 mutant differentially accumulates proanthocyanidin and anthocyanin through a 3′ amino acid substitution in glutathione S-transferase, Plant, Cell Environ. 34 (2011) 374–388.
[31] X. Liu, J. Huang, Y. Wang, K. Khanna, Z. Xie, H.A. Owen, D. Zhao, The role of floral organs in carpels, an Arabidopsis loss-of-function mutation in MicroRNA160a, in organogenesis and the mechanism regulating its expression, Plant J. 62 (2010) 416–428.
[32] K.J. Livak, T.D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2- ΔΔCT method, methods 25 (2001) 402–408.
[33] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 1–21.
[34] X. Ma, F. Zhao, B. Zhou, The characters of non-coding RNAs and their biological roles in plant development and abiotic stress response, Int. J. Mol. Sci. 23 (2022) 4124.
[35] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBnet. J. 17 (2011) 10–12.
[36] S. Mathur, K. Paritosh, R. Tandon, D. Pental, A.K. Pradhan, Comparative analysis of seed transcriptome and coexpression analysis reveal candidate genes for enhancing seed size/weight in *Brassica juncea.* Front. Genet. 13 (2022), 814486.
[37] S. Miyashima, M. Honda, K. Hashimoto, K. Tatematsu, T. Hashimoto, K. Sato-Nara, K. Okada, K. Nakajima, A comprehensive expression analysis of the Arabidopsis MICRORNA165/6 gene family during embryogenesis reveals a conserved role in meristem specification and a non-cell-autonomous function, Plant Cell Physiol. 54 (2013) 375–384.
[38] G. Na, X. Mu, P. Grabowski, J. Schmutz, C. Lu, Enhancing microRNA 167A expression in seed decreases the α-linolenic acid content and increases seed size in *Camelina sativa*, Plant J. 98 (2019) 346–358.
[39] Y. Ni, H.-L. Jiang, B. Lei, J.-N. Li, Y.-R. Chai, Molecular cloning, characterization and expression of two rapeseed (*Brassica napus L.*) cDNAs orthologous to *Arabidopsis thaliana phenylalanine ammonia-lyase 1*, Euphytica 159 (2008) 1–16.
[40] L. Ning, Z. Lin, J. Gu, L. Gan, Y. Li, H. Wang, L. Miao, L. Zhang, B. Wang, M. Li, The initial deficiency of protein processing and flavonoids biosynthesis were the main mechanisms for the male sterility induced by SX-1 in *Brassica napus*, BMC Genom. 19 (1) (2018) 18.
[41] Y. Niu, L. Wu, Y. Li, H. Huang, M. Qian, W. Sun, H. Zhu, Y. Xu, Y. Fan, U. Mahmood, Deciphering the transcriptional regulatory networks that control size, color, and oil content in *Brassica rapa* seeds, Biotechnol. Biofuels 13 (2020) 1–20.
[42] M.D. Nodine, D.P. Bartel, MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis, Genes Dev. 24 (2010) 2678.
[43] M. Nozawa, S. Miura, M. Nei, Origins and evolution of microRNA genes in plant species, Genome Biol. Evol. 4 (2012) 230–239.
[44] M. Ohto, S.K. Floyd, R.L. Fischer, R.B. Goldberg, J.J. Harada, Effects of APETALA2 on embryo, endosperm, and seed coat development determine seed size in Arabidopsis, Sex. Plant Reprod. 22 (2009) 277–289.
[45] J. Pan, D. Huang, Z. Guo, Z. Kuang, H. Zhang, X. Xie, Z. Ma, S. Gao, M.T. Lerdau, C. Chu, Overexpression of microRNA408 enhances photosynthesis, growth, and seed yield in diverse plants, J. Integr. Plant Biol. 60 (2018) 323–340.
[46] K. Paritosh, S.K. Yadava, P. Singh, L. Bhayana, A. Mukhopadhyay, V. Gupta, N. C. Bisht, J. Zhang, D.A. Kudrna, D. Copetti, A chromosome-scale assembly of allotetraploid *Brassica juncea* (AABB) elucidates comparative architecture of the A and B genomes, Plant Biotechnol. J. 19 (2021) 602–614.
[47] E. Petrussa, E. Braidot, M. Zancani, C. Peresson, A. Bertolini, S. Patui, A. Vianello, Plant flavonoids—biosynthesis, transport and involvement in stress responses, Int. J. Mol. Sci. 14 (2013) 14950–14973.
[48] M.S. Pidkowich, H.T. Nguyen, I. Heilmann, T. Ischebeck, J. Shanklin, Modulating seed β-ketoacyl-acyl carrier protein synthase II level converts the composition of a temperate seed oil to that of a palm-like tropical oil, Proc. Natl. Acad. Sci. 104 (2007) 4742–4747.

[49] A. Rani, S. Singh, P. Yadav, H. Arora, I. Kaur, N. Dhaka, MicroRNAs for understanding and improving agronomic traits in oilseed Brassicas, Plant Gene (2023), 100422.
[50] K. Rout, B.G. Yadav, S.K. Yadava, A. Mukhopadhyay, V. Gupta, D. Pental, A. K. Pradhan, QTL landscape for oil content in *Brassica juncea*: analysis in multiple bi-parental populations in high and "0" erucic background, Front. Plant Sci. 9 (2018), 1448.
[51] J.V. Sagun, U.P. Yadav, A.P. Alonso, Progress in understanding and improving oil content and quality in seeds, Front. Plant Sci. 14 (2023) 1116894.
[52] C. Schommer, J.M. Debernardi, E.G. Bresso, R.E. Rodriguez, J.F. Palatnik, Repression of cell proliferation by miR319-regulated TCP4, Mol. Plant 7 (2014) 1533–1544.
[53] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504.
[54] D. Sharma, M. Tiwari, A. Pandey, C. Bhatia, A. Sharma, P.K. Trivedi, MicroRNA858 is a potential regulator of phenylpropanoid pathway and plant development, Plant Physiol. 171 (2016) 944–959.
[55] S. Shen, Y. Tang, C. Zhang, N. Yin, Y. Mao, F. Sun, S. Chen, R. Hu, X. Liu, G. Shang, Metabolite profiling and transcriptome analysis provide insight into seed coat color in *Brassica juncea*, Int. J. Mol. Sci. 22 (2021) 7215.
[56] J.B. Song, X.X. Shu, Q. Shen, B.W. Li, J. Song, Z.M. Yang, Altered fruit and seed development of transgenic rapeseed (*Brassica napus*) over-expressing microRNA394, PLoS One 10 (2015), e0125427.
[57] A. Srivastava, V. Gupta, D. Pental, A. Pradhan, AFLP-based genetic diversity assessment amongst agronomically important natural and some newly synthesized lines of *Brassica juncea*, Theor. Appl. Genet. 102 (2001) 193–199.
[58] M. Tan, J. Niu, D.Z. Peng, Q. Cheng, M.B. Luan, Z.Q. Zhang, Clone and function verification of the OPR gene in *Brassica napus* related to linoleic acid synthesis, BMC Plant Biol. 22 (1) (2022) 19.
[59] M. Tang, X. Bai, L.-J. Niu, X. Chai, M.-S. Chen, Z.-F. Xu, miR172 regulates both vegetative and reproductive development in the perennial woody plant *Jatropha curcas*, Plant Cell Physiol. 59, 2549–2563 (2018).
[60] F. Tian, D.-C. Yang, Y.-Q. Meng, J. Jin, G. Gao, PlantRegMap: charting functional regulatory maps in plants, Nucleic Acids Res. 48 (2020) D1104–D1113.
[61] V. Tirumalai, C. Swetha, A. Nair, A. Pandit, P.V. Shivaprasad, miR828 and miR858 regulate VvMYB114 to promote anthocyanin and flavonol accumulation in grapes, J. Exp. Bot. 70 (2019) 4775–4792.
[62] B. Usadel, F. Poree, A. Nagel, M. Lohse, A. Czedik-Eysenberg, M. Stitt, A guide to using MapMan to visualize and compare Omics data in plants: a case study in the crop species, Maize, Plant, Cell Environ. 32 (2009) 1211–1229.
[63] R. Voorrips, MapChart: software for the graphical presentation of linkage maps and QTLs, J. Hered. 93 (2002) 77–78.
[64] J. Wang, H. Jian, T. Wang, L. Wei, J. Li, C. Li, L. Liu, Identification of microRNAs actively involved in fatty acid biosynthesis in developing *Brassica napus* seeds using high-throughput sequencing, Front. Plant Sci. 7 (2016) 1570.
[65] X. Wang, Y. Long, Y. Yin, C. Zhang, L. Gan, L. Liu, L. Yu, J. Meng, M. Li, New insights into the genetic networks affecting seed fatty acid concentrations in *Brassica napus*, BMC Plant Biol. 15 (2015) 1–18.
[66] Z. Wang, Y. Qiao, Jingjing Zhang, W. Shi, Jinwen Zhang, Genome wide identification of microRNAs involved in fatty acid and lipid metabolism of *Brassica napus* by small RNA and degradome sequencing, Gene 619 (2017) 61–70.
[67] W. Wei, G. Li, X. Jiang, Y. Wang, Z. Ma, Z. Niu, Z. Wang, X. Geng, Small RNA and degradome profiling involved in seed development and oil synthesis of *Brassica napus*, PloS One 13 (2018), e0204998.
[68] J. Xu, J. Li, L. Cui, T. Zhang, Z. Wu, P.-Y. Zhu, Y.-J. Meng, K.-J. Zhang, X.-Q. Yu, Q.-F. Lou, New insights into the roles of cucumber TIR1 homologs and miR393 in regulating fruit/seed set development and leaf morphogenesis, BMC Plant Biol. 17 (2017) 1–14.
[69] W. Yang, G. Wang, J. Li, P.D. Bates, X. Wang, D.K. Allen, Phospholipase Dζ enhances diacylglycerol flux into triacylglycerol, Plant Physiol. 174 (2017) 110–123.
[70] X. Yao, J. Chen, J. Zhou, H. Yu, C. Ge, M. Zhang, X. Gao, X. Dai, Z.-N. Yang, Y. Zhao, An essential role for miRNA167 in maternal control of embryonic and seed development, Plant Physiol. 180 (2019) 453–464.
[71] J. Zhan, B.C. Meyers, Plant Small RNAs: Their Biogenesis, Regulatory Roles, and Functions, Annu. Rev. Plant Biol. 74 (2023) 21–51.
[72] B. Zhang, X. Pan, S. Cox, G. Cobb, T. Anderson, Evidence that miRNAs are different from other RNAs, Cell. Mol. Life Sci. CMLS 63 (2006) 246–254.
[73] L. Zhang, Y. Xiang, S. Chen, M. Shi, X. Jiang, Z. He, S. Gao, Mechanisms of microRNA biogenesis and stability control in plants, Front. Plant Sci. 13 (2022), 844149.
[74] L. Zhang, B. Yang, C. Zhang, H. Chen, J. Xu, C. Qu, K. Lu, J. Li, Genome-wide identification and posttranscriptional regulation analyses elucidate roles of key argonautes and their miRNA triggers in regulating complex yield traits in rapeseed, Int. J. Mol. Sci. 24 (2023) 2543.
[75] M. Zhang, H. Zheng, L. Jin, L. Xing, J. Zou, L. Zhang, C. Liu, J. Chu, M. Xu, L. Wang, miR169o and *ZmNF-YA13* act in concert to coordinate the expression of *ZmYUC1* that determines seed size and weight in maize kernels, N. Phytol. 235 (2022) 2270–2284.
[76] Z. Zhang, J.M. Dunwell, Y.-M. Zhang, An integrated omics analysis reveals molecular mechanisms that are associated with differences in seed oil content between *Glycine max* and *Brassica napus*, BMC Plant Biol. 18 (2018) 1–15.
[77] H. Zhao, U. Basu, B. Kebede, C. Qu, J. Li, H. Rahman, Fine mapping of the major QTL for seed coat color in *Brassica rapa* var. Yellow Sarson by use of NIL populations and transcriptome sequencing for identification of the candidate genes, PLoS One 14 (2019), e0209982.
[78] Y.-T. Zhao, M. Wang, S.-X. Fu, W.-C. Yang, C.-K. Qi, X.-J. Wang, Small RNA profiling in two *Brassica napus* cultivars identifies microRNAs with oil production- and development-correlated expression and new small RNA classes, Plant Physiol. 158 (2012) 813–823.
[79] L. Zheng, X. Zhang, H. Zhang, Y. Gu, X. Huang, H. Huang, H. Liu, J. Zhang, Y. Hu, Y. Li, The miR164-dependent regulatory pathway in developing maize seed, Mol. Genet. Genom. 294 (2019) 501–517.
[80] Q.-H. Zhu, N.M. Upadhyaya, F. Gubler, C.A. Helliwell, Over-expression of miR172 causes loss of spikelet determinacy and floral organ abnormalities in rice (*Oryza sativa*), BMC Plant Biol. 9 (1) (2009) 13.