Full Length Article

Identification and Characterization of MicroRNAs during Common Buckwheat (*Fagopyrum esculentum*) Seed Development

Hongyou Li[1*], Xiaoqian Sun[1], Chao Ma[2], Hengling Meng[3] and Qingfu Chen[1]

[1]*Research Center of Buckwheat Industry Technology, Guizhou Normal University, Guiyang 550001, P.R. China*
[2]*College of Agriculture, Henan University of Science and Technology, Luoyang 471023, P.R. China*
[3]*Key Laboratory of High-Quality Crops Cultivation and Safety Control of Yunnan Province, Honghe University, Honghe, 661100, P.R. China*
[*]For correspondence: lihongyouluod@163.com

Abstract

MicroRNAs (miRNAs) have vital regulatory roles in many plant developmental processes, including seed development. Although a numerous of miRNAs had been characterized in many plants, the miRNAs in common buckwheat had not been discovered. In this study, the small RNAs profiles from three developmental stages common buckwheat seeds were gained by high-throughput sequencing. A total of 247 miRNAs, including 96 conserved and 151 novel miRNAs, were firstly identified in common buckwheat. For among 242 miRNAs, 15,403 potential target genes were identified. Expression profiles analysis found that 49 miRNAs (31 conserved miRNAs and 18 novel miRNAs) showed differential expression during common buckwheat seed development, and 8 among them were verified by qRT-PCR. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of the corresponding target genes of these differentially expressed miRNAs showed that the differentially expressed miRNAs participated in cell division, plant hormone signal transduction, and multiple metabolic processes. 4 miRNAs *i.e.*, fes-miR156j-5p, fes-miR396a-5p, fes-miR530b-5p, and fes_novel_miR48, were determine as the critical candidate miRNAs for regulating seed size according to NR functional annotation and homologous query of target genes. In addition, the expression negative correlation and cleavage sites between 4 miRNAs and their corresponding target TFs were confirmed by qRT-PCR and 5′ RLM-RACE, respectively. Our study provided the firsthand miRNAs data in common buckwheat, and insights into the dynamic expression profiles of miRNAs during common buckwheat seed development. © 2020 Friends Science Publishers

Keywords: Common buckwheat; Seed development; High-throughput sequencing; microRNAs; Target genes

Introduction

MicroRNAs (miRNAs) are a class of short endogenous non-coding small RNAs (sRNAs) with a length of 20–24 nucleotides (nt) in eukaryotic organisms (Tian *et al*. 2014). They had been generally recognized as crucial regulators of gene expression through negatively regulating their target genes expression at transcriptional or post-transcriptional levels (Kang *et al*. 2014; Li *et al*. 2016a; Ma *et al*. 2019). In plants, a large number of studies had indicated that miRNAs had vital roles in many biological processes including the development of roots, shoots, leaves and flowers, secondary substance metabolism, biotic and abiotic stress tolerance and so on (Bulgakov and Avramenko 2015; D'Ario *et al*. 2017; Ma *et al*. 2019; Hu *et al*. 2019; Zhang *et al*. 2019).

Seeds of crop plants were not only a unit of reproduction, but also as the primary storage organs for nutrients such as starch, lipids, and proteins (Huang *et al*.

2017). Therefore, the successful development of seeds would directly determine seed crop yield and seed quality and the successful racial continuation of crop plants. Plant seed development was an extremely complex and dynamic process, which involved in the expression regulation of numerous genes at transcriptional and post-transcriptional levels (Samad *et al*. 2017; Savadi 2017; Li *et al*. 2019a). Currently, the increasing body of evidence showed that miRNAs played a key regulatory role in plant seed development. Using high-throughput sequencing approach, thousands of miRNAs had been identified in the developing seed of multiple crop plants, of which many miRNAs were specially or differentially expressed in the developing seed, indicating the significance of miRNAs in seed development (Li *et al*. 2015; Wang *et al*. 2015, 2016; Bai *et al*. 2017; Wei *et al*. 2018; Yu *et al*. 2019; DeBoer *et al*. 2019). In fact, some miRNAs had been functionally demonstrated to regulate the seed development, with focus on in rice. For

example, OsmiR397 and OsmiR408 positively regulated the grain size and weight by down-regulating their targets *OsLAC* and *OsUCL8* expression (Zhang *et al*. 2013, 2017a), respectively. Interestingly, over-expression miR408 of *Arabidopsis* and tobacco also increased the seed size, suggesting that miR408 had a conserved function in the determination of seed size in seed plants (Pan *et al*. 2018). Furthermore, OsmiR398 and OsmiR159 were also revealed that positively regulated rice grain size using short tandem target mimic (STTM) technology (Zhang *et al*. 2017b; Peng *et al*. 2018). On the contrary, OsmiR156 and OsmiR396 targeted *OsSPL14* and *OsGRF4*, respectively, and negatively regulated the grain size (Jiao *et al*. 2010; Duan *et al*. 2015; Li *et al*. 2016b). Likely, OsmiR1432 and OsmiR167 were also demonstrated to play a negative regulatory role in grain size using STTM technology (Peng *et al*. 2018; Zhao *et al*. 2019). In addition, a few of other miRNAs were demonstrated to regulate the nutrients accumulation during seed development. For example, *OsmiR160* positively regulated the starch accumulation and affected grain size through negative regulating its target *OsARF18* expression (Huang *et al*. 2016). Overexpression EgmiR5179 of oil palm (*Elaeis guineensis* Jacq.) in *Arabidopsis* increased the content of seeds oil and linolenic acid (Gao *et al*. 2019), indicating that EgmiR5179 played a positive regulatory role in oil and linolenic acid accumulation. Similarly, in *Camelina sativa*, CsmiR167A had positive and negative effects on the biosynthesis of linolenic acid and α-linolenic acid, respectively (Na *et al*. 2019).

Common buckwheat (*Fagopyrum esculentum*) is an important pseudocereal crop that widely cultivated in the mountainous areas of Asia due to its rich seed nutrients such as starch, protein, fatty acid, vitamins, minerals, flavonoids, and dietary fibers (Yasui *et al*. 2016; Li *et al*. 2019b). Therefore, it is very important to insight into the molecular mechanisms of common buckwheat seed development, as it will contribute to the seed improvement of common buckwheat especially nutrient accumulation and seed size. So far, several studies have been performed to investigate the gene expression profile of the developing common buckwheat seeds using RNA-Seq (Gao *et al*. 2017; Shi *et al*. 2017; Li *et al*. 2019b). However, as far as we know, miRNAs in common buckwheat have not been identified and characterized, and miRNA-mediated regulation in developing common buckwheat seeds needs to be explored.

In this study, genome-wide identification of miRNAs was conducted in developing common buckwheat seeds using high-throughput sequencing technology. Through comprehensive analysis, the conserved, new, and differentially expressed miRNAs were identified. Furthermore, the differentially expressed miRNAs and its target genes were analyzed, and the candidate miRNAs that involved in seed development including plant hormone signal transduction and seed size were also identified based on their target genes annotation. These results extend our

knowledge for miRNAs participated in the seed development of common buckwheat.

Materials and Methods

Plant materials and samples collection

Common buckwheat cultivar "Chitian No. 1" was planted in the experimental field of the Guizhou Normal University (Guiyang, China; Lat. 26°62' N, 106°72' E, Alt. 1168 m) on 20 March 2018. The field management, seed samples collection was performed as previously described (Li *et al*. 2019b). Seeds were harvested at the 8th, 14th and 21st days after pollination (Li *et al*. 2019b).

Construction and sequencing of small RNA libraries

Total RNA was isolated from three samples using the TaKaRa MiniBEST Plant RNA Extraction Kit (TaKaRa, Dalian, China) based on the manufacturer's instructions. RNA samples quality was examined with a NanoDrop 2000c spectrophotometer (NanoDrop, Wilmington, DE, USA) and 1.2% agarose gel electrophoresis, respectively. Small RNA libraries were constructed according to the previous description by (Xu *et al*. 2018) and subjected to high-throughput sequencing using Illumina SE50 system at Biomarker Technologies Co., Ltd. (Beijing, China).

Small RNA (sRNA) sequencing data analysis

Clean reads were generated by removing adaptor sequences and low-quality reads in raw data. Then, the obtained clean reads were used to search on Silva database, GtRNAdb database, Rfam 11.0 database and Repbase database to filter the rRNA, tRNA, snRNA and snoRNA to product the unannotated clean reads that containing miRNA using the Bowtie software according to default parameters (Langmead *et al*. 2009). The unannotated clean reads were further mapped on the common buckwheat reference genome (ftp://ftp.kazusa.or.jp/pub/buckwheat/) (Yasui *et al*. 2016; Li *et al*. 2019b) using the Bowtie software. After that, these perfectly matched clean reads were blasted against in miRBase v. 21.0 (http://www.mirbase.org/index.shtml) to identify the known miRNAs, and the rest unmapped sequences were used to predict novel miRNAs using miRdeep2 program with score ≤ 5 (Zhang *et al*. 2015).

Identification of differential expression miRNAs during common buckwheat seed development

To screen differentially expressed miRNAs in the course of common buckwheat seed development, the expression level of miRNAs in each sample was normalized to TPM using miRNA counts. Then, the significantly differential expression miRNAs were identified by using the IDEG6 software with $|log2(fold change)| \geq 1$ and $P \leq 0.05$ as the threshold (Romualdi *et al*. 2003).

Prediction and analysis of miRNA target genes

The TargetFinder software was used to predict the potential miRNAs target genes with default parameters (Allen *et al.* 2005). Furthermore, the predicted target genes were further function annotated in Blast2GO 5.0 software (http://www.blast2go.org/), and search on the KEGG pathway database to elucidate the pathways, respectively.

qRT-PCR analysis of miRNAs and their targets

To verify the sequencing data, 12 miRNAs and 4 corresponding target TFs of among 4 miRNAs were selected to perform qRT-PCR analysis. The RNA samples were used for qRT-PCR from these samples that used for small RNA libraries construction. RT-qPCR was conducted on a CFX96 Real-time System (BIO-RAD, U.S.A.) using SYBR® Premix Ex TaqTM II Kit (TaKaRa, Dalian, China) and Mir-X miRNA qRT-PCR TB Green® Kit (TaKaRa, Dalian, China), respectively. For miRNAs, the mature miRNA sequences were used to design the forward primers and the adapter sequences from the cDNA synthesis kit (TaKaRa, Dalian, China) were used as the reverse primers. The *U6* snRNA and *Actin7* of common buckwheat were used as the references for miRNAs and their target genes, respectively. Three biological replicates with three technical replicates were performed for each sample. The relative expression level of each miRNA or target gene was calculated by the $2^{-\Delta\Delta Ct}$ method. All primers used for qRT-PCR were listed in Table S1.

RLM-5′ RACE

The 5′-RLM-RACE RNA method was used to examine the miRNA-directed cleavage for their predicted target genes in vivo. The 5′-RLM-RACE was carried out as described previously (DeBoer *et al.* 2019). The primers were listed in Table S2.

Results

sRNA sequencing from developing common buckwheat seeds

Nine sRNA libraries derived from three different developmental stages common buckwheat seeds were constructed and sequenced. Upon sequencing, a total of 103,446,033, 57,491,562, and 54,022,652 reads, on average, were obtained for each sample, respectively (Table 1). After removing the adaptor sequences, low-quality reads and reads with lengths < 18 or > 30, a total of 10,412,548 to 44,809,759 clean reads were obtained for each library (Table 1). As shown in Fig. 1, most sRNA sequences in the nine libraries were 21–24 nt long. Among them, 24 nt sRNAs were the most abundant sRNAs in all samples, followed by 23 and 21 nt. In addition, the number of 21–24 nt sRNAs was dynamic change during the seed development, with a decreased trend (Fig. 1). Among these clean reads, 16.93, 16.11, 17.18, 30.85, 21.30, 26.67, 52.28, 55.48 and 51.07% reads for each library were annotated as rRNA, tRNA, snRNA, snoRNA, and Repbase (Table 1). Furthermore, a total of 49.15, 49.22, 50.01, 46.75, 45.89, 47.23, 49.41, 48.59 and 47.34% unannotated sRNA for each library were mapped to the common buckwheat reference genome (Table 1).

Conserved miRNAs in developing common buckwheat seeds

A total of 96 conserved miRNAs, which divided into 24 families, were identified (Table S3). Among these families, 3 families (miR156, miR166, and miR319) contained ≥ 10 members, 17 families contained 2 to 7 members, and 4 families (miR159, miR394, miR828, and miR858) only contained 1 member (Fig. 2a). The lengths of these conserved miRNAs were 19 to 22 nt, of which 22 nt contained the largest number of miRNAs (52) (Fig. 2b). Among these conserved miRNAs, about 60% miRNAs (57) were lowly expressed with < 100 read counts in each library (Table S3). Other 17 and 22 miRNAs were expressed at a high (read counts > 1000 at least in one sample) and moderate level (100 ≤ read counts ≤ 1000) (Table S3), respectively. The fes-miR159-3p had the highest expression in all 96 conserved miRNAs, with the read counts ranging from 12,912 to 48,871 for each library (Table S3).

Novel miRNAs in developing common buckwheat seeds

To further identify novel miRNAs in common buckwheat seeds, miRdeep2 program was utilized. As a result, a total of 151 novel miRNAs were predicted (Table S4). Of these, 57 miRNAs were divided into 36 currently known miRNA families while the remained 94 miRNAs did not show similarity with any known family in the miRNA database. The lengths of these novel miRNAs were 18 nt (1), 20 nt (1), 21 nt (53), 22 nt (19), 23 nt (6), and 24 nt (71), and the minimum free energy (MFE) distributed between -113.2 kcal moL-1 to -10.6 kcal moL^{-1} (Table S4). In addition, the precursor lengths of these novel miRNAs ranged from 64 to 250 nt (Table S4). Among these novel miRNAs, most miRNAs were moderately or highly expressed in the developmental common buckwheat seeds. The fes_novel_miR48 displayed the highest expression with the read counts ranging from 46,207 to 78,975 for each library (Table S4).

Prediction and functional annotation of the miRNAs targets

Target prediction is a good way to understand the functions of miRNAs. Using the TargetFinder software, 15,403 potential target genes of 242 miRNAs were successfully

Table 1: Statistical analysis of sequencing reads in nine sRNA libraries in common buckwheat seeds

Type	S1-1	S1-2	S1-3	S2-1	S2-2	S2-3	S3-1	S3-2	S3-3
Raw reads	26,403,775	22,649,461	28,392,797	18,955,805	16,836,019	21,699,738	15,928,753	18,669,588	19,424,311
Clean reads	21,545,480	18,796,381	22,932,862	14,060,697	13,617,617	16,986,555	10,412,548	10,623,054	12,787,024
	(81.6%)	(82.75%)	(80.77%)	(72.49%)	(74.18%)	(78.28%)	(67.34%)	(61.95%)	(65.83%)
Q30 (%)	98.17	98.23	98.20	98.23	98.28	98.35	98.23	98.33	98.25
rRNA	3,348,167	2,681,862	3,423,876	3,927,086	2,640,161	3,990,142	5,093,371	5,433,589	6,003,508
	(15.54%)	(14.27%)	(14.93%)	(27.93%)	(19.39%)	(23.49%)	(48.92%)	(51.15%)	(46.95%)
tRNA	288,709	339,167	506.816	405,597	253,042	531,679	344,843	455,670	521,711
	(1.34%)	(1.80%)	(2.21%)	(2.88%)	(1.86%)	(3.13%)	(3.31%)	(4.29%)	(4.08%)
snRNA	0 (0.00%)	1 (0.00%)	0	0	1	1	0	0	0
scRNA	0 (0.00%)	0 (0.00%)	0	0	0	0	0	0	0
snoRNA	2,136	1,608	2,093	763	918	1,547	564	364	636
	(0.01%)	(0.01%)	(0.01%)	(0.01%)	(0.01%)	(0.01%)	(0.01%)	(0.00%)	(0.01%)
Repbasea	8,618	6,375	6,880	4,697	5,954	6,794	3,876	4,053	3836
	(0.04%)	(0.03%)	(0.03%)	(0.03%)	(0.04%)	(0.04%)	(0.04%)	(0.04%)	(0.03%)
Unannotated	17,897,850	15767368	18,993,196	9,722,554	10,717,541	12,456,392	4,969,894	4,729,378	6,257,333
	(83.07%)	(83.89%)	(82.82%)	(69.15%)	(78.70%)	(73.33%)	(47.72%)	(44.52%)	(48.93%)
Mapped reads	8,796,793	7,760,203	9,498,497	4,544,911	4,918,807	5,883,154	2,455,549	2,298,005	2,962,221
	(49.15%)	(49.22%)	(50.01%)	(46.75%)	(45.89%)	(47.23%)	(49.41%)	(48.59%)	(47.34%)

Notes: S1, S2, and S3 represent the seed samples at three different developmental stages (pre-filling stage, filling stage, and initial maturity stage), and -1, -2, and -3 stand for three replicates of each sample
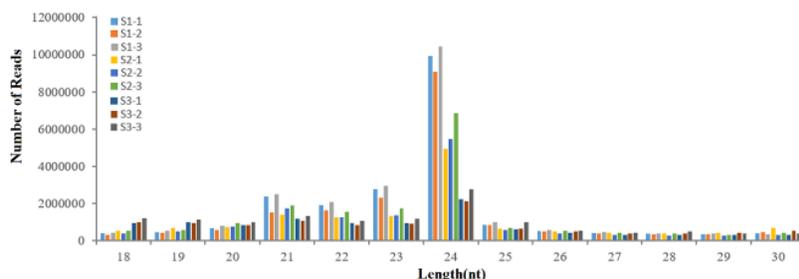

Fig. 1: Read length distribution of small RNAs. S1, S2, and S3 represent the seed samples at three different developmental stages (pre-filling stage, filling stage, and initial maturity stage), and -1, -2, and -3 stand for three replicates of each sample

predicted. Of which, 96 conserved miRNAs targeted 8997 genes, and 146 novel miRNAs targeted 6862 genes. All predicted target genes were further annotated in eight databases. As a result, 3,348, 4,305, 1,209, 7,879, 7,760, 5,674, 9,861, and 10552 target genes were annotated in CGO, GO, KEGG, KOG, Pfam, Swissprot, eggNOG, and NR databases, respectively. The target genes from GO annotations were further subjected functional classification (biological process, cellular component, and molecular function). The 4,305 GO annotated target genes were assigned to 967 unique biological processes, 188 unique cellular components, and 566 unique molecular functions (Table S5). The top 15 GO groups of these three categories were represented in Fig. 3. As shown in Fig. 3, a numerous of target genes involved in DNA metabolic process (GO:0006259), DNA integration (GO:0015074), nucleus (GO:0005634), membrane (GO:0016020), binding (GO:0005488), and nucleic acid binding (GO:0003676) (Fig. 3 and Table S5). In addition, based on the NR functional annotation results, 260 transcription factors (TFs) from 34 TF families were identified as the potential target genes of 141 miRNAs (including 79 conserved and 62 novel miRNAs (Table S6). The top five largest TF families were *MYB* (48), *MYB*-related (20), *FAR1* (18), *C2H2* (15),

GRAS (15), *bHLH* (13), *NAC* (13), and *SBP* (13) (Table S6).

Differentially expressed miRNAs during common buckwheat seed development

The expression abundance of these identified miRNAs were normalized to TPM. A good correlation existed between any two biological replicates of each sample (Table S7). Based on a pairwise comparison of three samples, a total of 49 miRNAs shown significantly differentially expressed during common buckwheat seed development (Fig. 4 and Table S8). When compared S1 to S2, 25 miRNAs, including 10 up-regulated (7 conserved and 3 novel miRNAs) and 15 down-regulated (13 conserved and 2 novel miRNAs), were identified (Fig. 5 and Table S8). In comparison S2 vs S3, 9 miRNAs (2 conserved and 7 novel miRNAs) and 6 conserved miRNAs were up- or down-regulated, respectively (Fig. 5 and Table S8). For S1 vs S3 comparison, 29 miRNAs were differentially expressed with 13 up-regulated (4 conserved and 9 novel miRNAs) and 16 down-regulated (11 conserved and 5 novel miRNAs) (Fig. 5 and Table S8). In addition, 8 (fes-miR166b-3p, fes-miR166g-3p, fes-miR395a-3p, fes-miR395b-3p, fes-miR396f-3p, fes-miR398c-3p, fes-
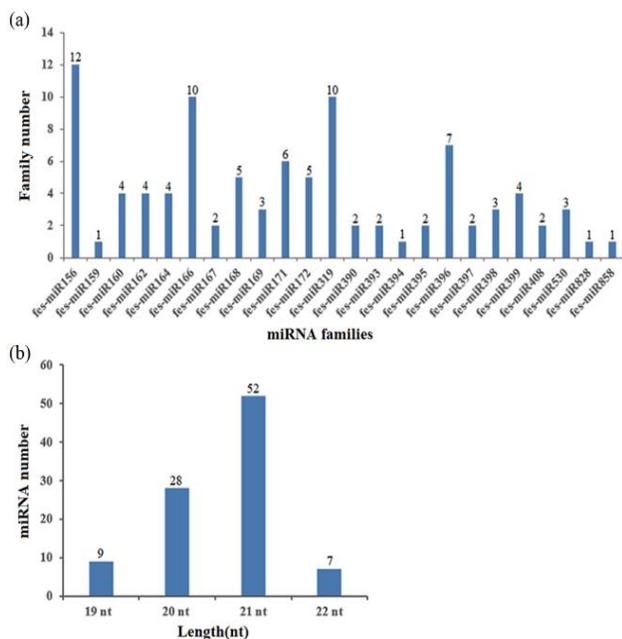
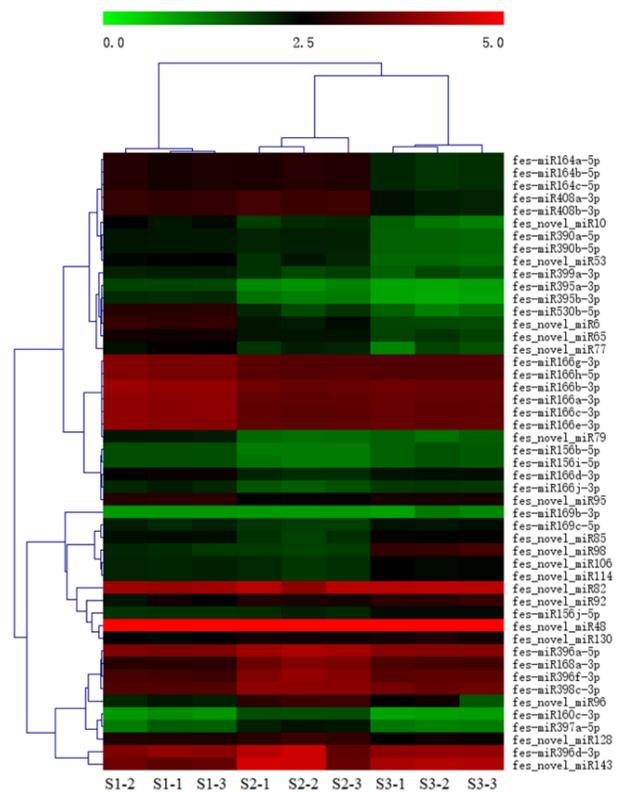**Fig. 2:** The number of miRNA family members (**a**) and the length distribution of conserved miRNAs (**b**)
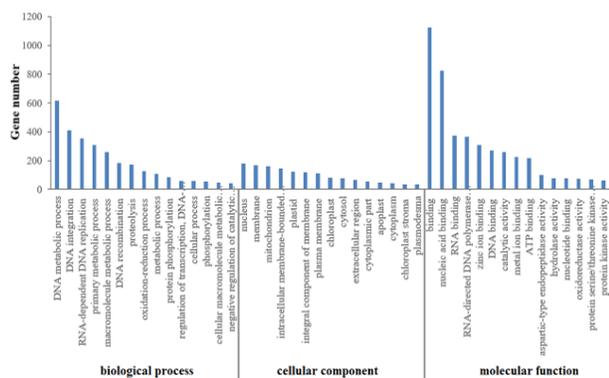
Fig. 3: GO annotations of the predicted target genes of identified miRNAs in common buckwheat seed
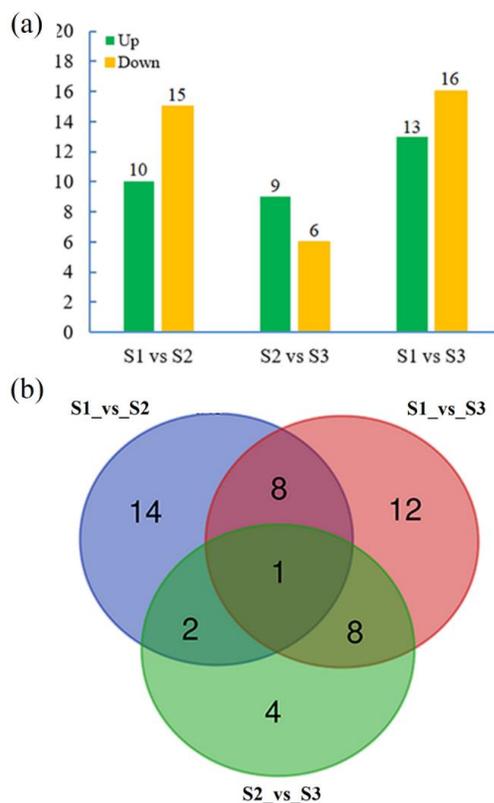

Fig. 4: Heat map diagrams of the differentially expressed miRNAs in the different developmental stages of common buckwheat seed. MeV 4.9.0 software was used to construct the heat map based on the $Log_2(TPM)$ value of miRNAs. S1, S2, and S3 represent the seed samples at three different developmental stages (pre-filling stage, filling stage, and initial maturity stage), and -1, -2, and -3 stand for three replicates of each sample
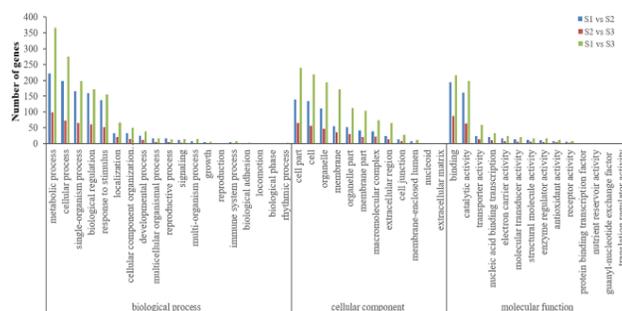
Functional prediction of the differentially expressed miRNAs

To provide insight into the potential regulatory functions of these differentially expressed miRNAs, target genes of them were searched in the total predicted target genes. As a result, these 49 differentially expressed miRNAs were found that have 5845 potential target genes (Table S9). Of these, fes_novel_miR95 and fes_novel_miR143 had the smallest number (1) and the largest number (937) target genes, respectively. All the other miRNAs had multiple target genes. In addition, the same target gene could be targeted by multiple miRNAs (Table S9).

Based on the GO enrichment result, these target genes were enriched into 36 functional groups, including 19 biological process categories, 12 cellular component categories, and 14 molecular function categories (Fig. 6). For biological process category, the 'metabolic process', 'cellular process' and 'single-organism process' have the top three gene numbers (Fig. 6). Within cellular component category, the 'cell part', 'cell', and 'organelle' were the

miR530b-5p, fes_novel_miR6), 2 (fes-miR397a-5p and fes-miR166j-3p), and 8 (fes-miR164a-5p, fes-miR164b-5p, fes-miR164c-5p, fes_novel_miR85, fes_novel_miR98, fes_novel_miR106, fes_novel_miR114, fes_novel_miR130) miRNAs were existed both in S1 *vs.* S2 and S1 *vs.* S3, S1 *vs.* S2 and S2 *vs.* S3, and S2 *vs.* S3 and S1 *vs.* S3, respectively (Fig. 5 and Table S8). Lastly, 1 miRNA (fes_novel_miR48) contained in all three comparisons, which displayed up-trend during common buckwheat seed development (Fig. 5 and Table S8). Notably, the expressions of most miRNAs were changed in only one stage while the expressions of fes-miR166j-3p, fes-miR397a-5p, and fes_novel_miR48 were changed in all developmental stages (Fig. 4 and Table S8).

Fig. 5: Numbers of differentially expressed miRNAs in different comparison groups during common buckwheat seed development. (**a**) Numbers of up- and down regulated miRNAs in different comparison groups. (**b**) Venn diagram for differentially expressed miRNAs in different comparison groups. S1, S2, and S3 stand for the seed samples at pre-filling stage, filling stage, and initial maturity stage, respectively.

Fig. 6: GO enrichment of the differentially expressed miRNAs target genes

greatest abundance terms (Fig. 6). Under the molecular function category, 'binding', 'catalytic activity', and 'transporter activity' represented the terms with the highest gene numbers (Fig. 6).

KEGG pathway analysis revealed 141 target genes were significantly enriched into 44 pathways (Table S9). All these enriched pathways could be further classified into the organismal systems category, metabolism category, cellular

process category, environmental information processing category, and genetic information processing category. The metabolism category possessed the maximum number of enriched pathways (22) and target genes (67), of which pathways major involved in phenylpropanoid biosynthesis (ko00940), amino sugar and nucleotide sugar metabolism (ko00520), amino acids biosynthesis (ko01230), and purine metabolism (ko00230). Furthermore, in environmental information processing category, 8 target genes from 7 miRNAs were significantly enriched in the "plant hormone signal transduction" pathway (Table S9). In this pathway, four miRNAs, fes-miR160c-3p, fes_novel_miR96, fes_novel_miR130, and fes_novel_miR143, targeted *ARF*, *AUX*, *IAA*, and *SAUR* genes, respectively, to respond to the indole-3-acetic acid (IAA) signal (Fig. 7). Two miRNAs, fes-miR397a-5p and fes-miR164a-5p were involved in abscisic acid (ABA) and ethylene (ET) pathway though targeting *SnRK2* and *CTR1* genes (Fig. 7), respectively. Three miRNAs, fes-miR390a-5p, fes-miR390b-5p, and fes-miR156j-5p targeted *BRI1* and *CYCD3* genes, respectively, and were involved in brassinosteroid (BR) signal (Fig. 7).

In addition, based on the NR functional annotation result and homologous query of target genes, it was found that four miRNAs, fes-miR156j-5p, fes-miR396a-5p, fes-miR530b-5p, fes_novel_miR48, targeted the orthologs of 4 known seed size-related *AtIKU2*, *OsGRF4*, *OsGIF1*, and *AtANT* genes, respectively (Li *et al*. 2019a). These findings indicated that the four miRNAs might play a crucial role in control the common buckwheat seed size.

qRT-PCR validation of differentially expressed miRNAs

To validate the reliability of miRNA sequencing data, 8 differentially expressed miRNAs (4 conserved and 4 novel miRNAs) were further analyzed by qRT-PCR. As shown in Fig. 8a, except for fes_novel_miR85, all the examined miRNAs displayed similar expression patterns between the qRT-PCR results and sequencing data, indicating that the RNA-seq was accurate and reliable.
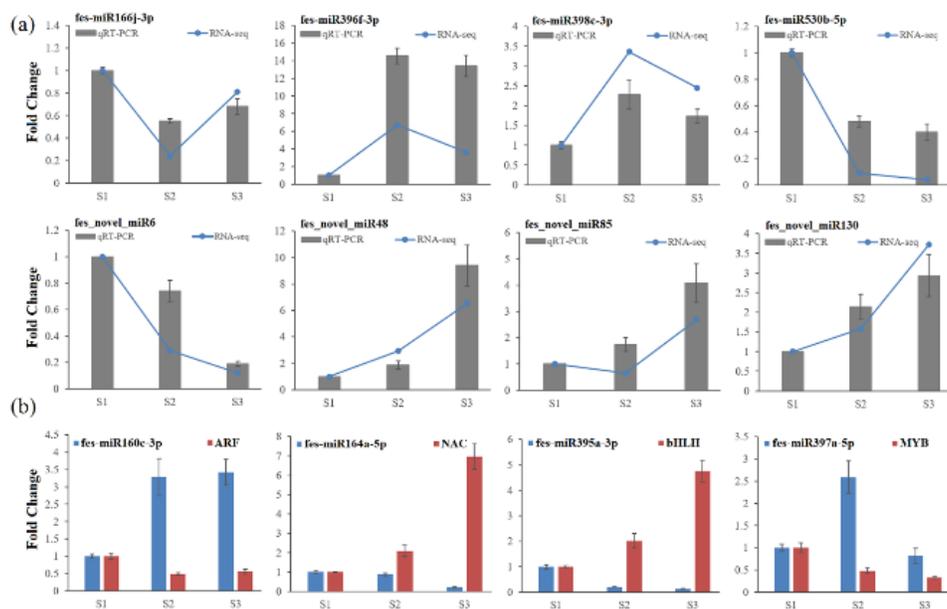
To further verify the expression relationship between miRNAs and their targets, 4 differentially expressed miRNAs (fes-miR160c-3p, fes-miR164a-5p, fes-miR395a-3p, fes-miR397a-5p) and their 4 predicted TF target genes (*ARF*, *NAC*, *bHLH*, and *MYB*) were subjected to qRT-PCR analyzed. As a result, the 4 miRNAs showed strong negative correlations with their potential target genes (Fig. 8b), indicating that these miRNAs may play important regulatory roles in common buckwheat seed development though negatively regulating these target TF expressions.

miRNA target validation

To further validate the cleavage sites in target transcripts, 4
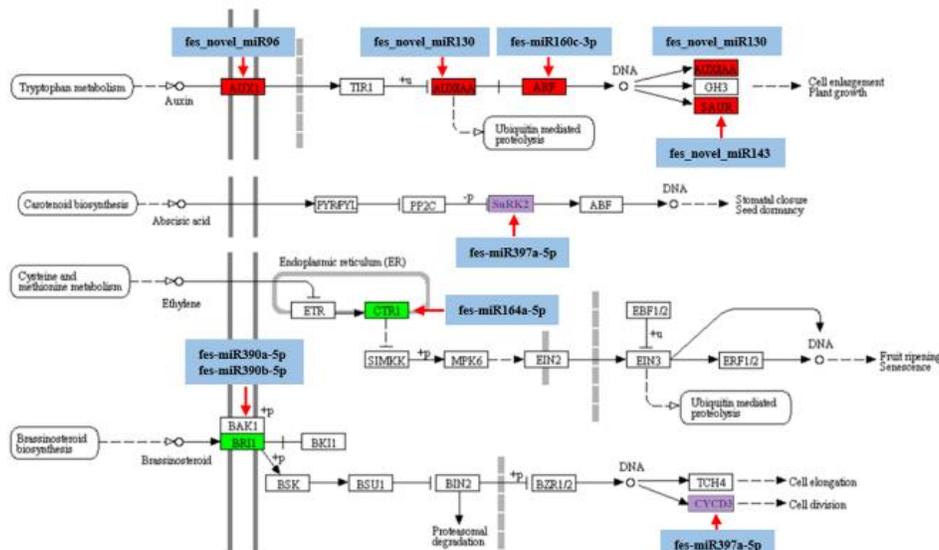
Fig. 7: KEGG pathways related to plant hormone signal transduction targeted by differentially expressed miRNAs.
Target genes labeled red represented the corresponding miRNAs were up-regulated; Target genes labeled green represented the corresponding miRNAs were down-regulated;
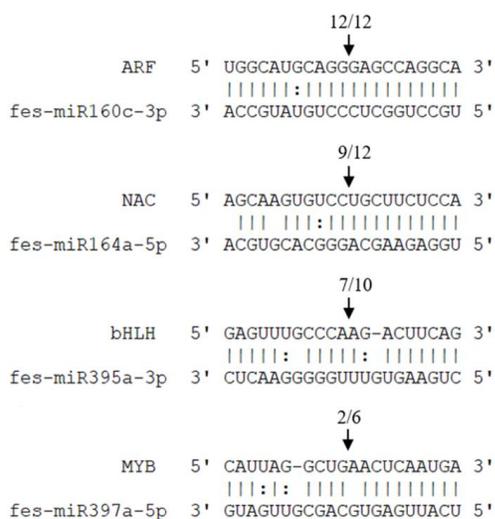Target genes labeled purple represented the corresponding miRNAs were up- or down-regulated

Fig. 8: Validation differentially expressed miRNAs and their target genes expression in developing common buckwheat seed using qRT-PCR. (**a**) Validation the expression of 8 differentially expressed miRNAs between small RNA sequencing and qRT-PCR. (**b**) Determination the expression of 4 differentially expressed miRNAs and their 4 target TFs during common buckwheat seed development. *U6* snRNA and *Actin7* as a standard for miRNA and target genes, respectively

target TFs (ARF, NAC, bHLH, and MYB) from previous qRT-PCR results were selected to perform 5′ RLM-RACE analysis. As results, 4 target TFs could be cleavage by 4 corresponding miRNAs (fes-miR160c-3p, fes-miR164a-5p, fes-miR395a-3p, fes-miR397a-5p), and the cleavage sites were mapped between the 10th and 11th nucleotide upstream of the miRNA 5′end (Fig. 9).

Discussion

Plant seed development was an extremely complex and dynamic process, which was involved in the expression regulation of numerous genes (Samad *et al*. 2017; Savadi 2017; Li *et al*. 2019a). Previous studies had indicated that miRNAs played crucial regulatory roles in seed

```
                        12/12
                          ↓
ARF            5' UGGCAUGCAGGGAGCCAGGCA 3'
                  ||||||:|||||||||||||
fes-miR160c-3p 3' ACCGUAUGUCCCUCGGUCCGU 5'

                        9/12
                          ↓
NAC            5' AGCAAGUGUCCUGCUUCUCCA 3'
                  ||| |||:|||||||||||
fes-miR164a-5p 3' ACGUGCACGGGACGAAGAGGU 5'

                        7/10
                          ↓
bHLH           5' GAGUUUGCCCAAG-ACUUCAG 3'
                  |||||: |||||: |||||||
fes-miR395a-3p 3' CUCAAGGGGGUUUGUGAAGUC 5'

                        2/6
                          ↓
MYB            5' CAUUAG-GCUGAACUCAAUGA 3'
                  |||:|: |||| ||||||||
fes-miR397a-5p 3' GUAGUUGCGACGUGAGUUACU 5'
```

Fig. 9: Mapping of the mRNA cleavage sites by 5′ RLM-RACE. Arrows indicated miRNAs cleavage sites. The numbers above sequences indicated the detected cleavage site of independent clones

development of many plants through negatively regulating its target genes expression (Wei *et al.* 2017). Nevertheless, miRNAs for seed development had not been well-explored in common buckwheat. In this study, we used the high-throughput sequencing technology to insight into the miRNA profiles and identified miRNAs that might be involved in common buckwheat seed development.

In total, 247 miRNAs (96 conserved and 151 novel) were first identified in common buckwheat. Of these, 242 miRNAs were predicted to target 15,403 genes. Among the 242 miRNAs, 141 miRNAs were found to target 260 TFs from 34 TF families, which were similar to previous reports that plant miRNAs perfected targeting transcription factors to perform potent functions in many developmental processes (Li and Zhang 2016). This suggested that the expression changes of TFs mediated by miRNAs were a conserved regulation network for different developmental processes in different species. According to bioinformatics analysis, we found that most conserved and novel miRNAs were lowly expressed, suggesting that they might have some roles in common buckwheat seed development. Notably, fes-miR159-3p was the highest expressed conserved miRNA in developing seeds at all the three developmental stages. In rice, miR159 was also found displayed higher expression in developing seed (Peng *et al.* 2014) and it was demonstrated that positively regulated grain length and width in rice (Jiao *et al.* 2010; Peng *et al.* 2018). Similarly, in tomato, miR159 also displayed higher expression in developing fruit and played a vital regulatory role in fruit development (Silva *et al.* 2017). These indicated that fes-miR159-3p might also have a key regulatory role in the development of tartary buckwheat seed, and miR159

regulated seed or fruit development was conserved in different plants. Similarly, fes_novel_miR48 was the highest expressed novel miRNA in developing common buckwheat seeds. Target gene prediction found that fes_novel_miR48 targeted four GRAS TFs, including two orthologs of *Arabidopsis SCL3* which as a direct target gene of DELLAs and positively responding to GA-signaling (Cui *et al.* 2007). Furthermore, in rice, a GRAS gene was demonstrated regulating seed development (Sun *et al.* 2013). These indicated that fes_novel_miR48 might also play a crucial regulatory role in common buckwheat seed development through GRAS mediating GA-signaling.

Generally, miRNAs expression pattern is closely related to its function. In multiple crop plants, many miRNAs had been found that differentially expressed in the developing seed (Li *et al.* 2015; Wang *et al.* 2015, 2016; Bai *et al.* 2017; Wei *et al.* 2018; Yu *et al.* 2019; DeBoer *et al.* 2019), implying they had key regulatory roles for seed development. In our study, 49 miRNAs (31 conserved and 18 novel), were identified as differentially expressed miRNAs during common buckwheat seed development. Notably, among the conserved miRNAs, most of them (miR156, miR160, miR164, miR166, miR168, miR169, miR390, miR395, miR396, miR397, miR398, miR399, and miR408) were also found that differentially expressed in other seed crops such as peanut (Ma *et al.* 2018), rice (Peng *et al.* 2014), wheat (Li *et al.* 2015), *Brassica napus* (Wei *et al.* 2018), maize (Li *et al.* 2016a). These findings indicated that these conserved miRNAs might have important and conserved regulatory roles in different plants seed development. It was well known that hormone signaling played a key role in seed development (Li *et al.* 2019a). In our study, based on KEGG enrichment analysis of the target genes of differentially expressed miRNAs, 4 miRNAs, including fes-miR160c-3p (targeting the *ARF*), fes_novel_miR96 (targeting *AUX1*), fes_novel_miR130 (targeting *IAA*), and fes_novel_miR143 (targeting *SAUR*), were found to be related to auxin signaling. In barley (Bai *et al.* 2017) and peanut (Ma *et al.* 2018), miR160 was also found to target ARF TFs gene and displayed increasing expression during seed development. Furthermore, in rice, *OsmiR160* had been demonstrated that positively regulated the grain development through negatively regulating its target *OsARF18* expression (Huang *et al.* 2016). In our study, we observed the expression level of fes-miR160c-3p and its target *ARF* were up- and down-regulated during common buckwheat seed development, respectively, through qRT-PCR analysis. These indicated that miR160 mediated *ARF* expression was a conserved auxin regulation mechanism in different plants seed development. In addition, fes-miR397a-5p (targeting *SnRK2*), fes-miR164a-5p (targeting *CTR1*), and fes-miR390a-5p (targeting *BRI1*), fes-miR390b-5p (targeting *BRI1*) and fes-miR156j-5 (targeting *CYCD3*) were found involved in ABA, ET, and BRs signaling. Together, these results suggested that miRNAs might have a significant effect on the regulation of

common buckwheat seed development by effecting different hormone signaling pathways. To date, there were over 100 genes had been demonstrated controlled the plant seed size (Li *et al*. 2019a). In our study, four miRNAs, fes-miR156j-5p, fes-miR396a-5p, fes-miR530b-5p, fes_novel_miR48, were found to target the orthologs of 4 known seed size-related *AtIKU2*, *OsGRF4*, *OsGIF1*, and *AtANT* genes, respectively. In *Arabidopsis*, *AtIKU2* had been demonstrated that positively regulated embryo and endosperm development (Luo *et al*. 2005). Interestingly, in our study, fes-miR156j-5p was lowly expressed in S1 and S2 periods (the key periods of embryo and endosperm development) and highly expressed in S3, suggesting that it might negatively regulate the embryo and endosperm development of common buckwheat seed through regulating its target gene *IKU2* expression. Notably, fes-miR396a-5p targeted the orthologs of *OsGRF4*. In rice, OsmiR396 was found to target *OsGRF4* and negatively regulated the grain size (Duan *et al*. 2015; Li *et al*. 2016b). This indicated that fes-miR396a-5p might also negatively regulate the seed size of common buckwheat, and miR396 might have a conserved role in regulating seed size in different plants. In addition, the expression negative correlation and cleavage sites between 4 differentially expressed miRNAs and their target TFs (ARF, NAC, bHLH and MYB) were confirmed by qRT-PCR and 5′ RLM-RACE, respectively, suggesting that these 4 miRNAs and their corresponding target TFs might play crucial regulation roles in common buckwheat seed development. Together, our results indicated that the complex regulatory mechanism of miRNAs for common buckwheat seed development. However, the detailed roles and mechanism underlying specific miRNAs involved in seed development remained to be investigated.

Conclusion

A total of 247 miRNAs, including 96 conserved and 151 novel miRNAs, were first identified in common buckwheat. 15,403 target genes were predicted, 8,997 for 96 conserved miRNAs and 6862 for146 novel miRNAs. 49 miRNAs, including 31 conserved and 18 novel miRNAs, displayed significantly up- or down-regulated during common buckwheat seed development. Some differentially expressed miRNAs that involved in hormone signaling and seed size were identified based on KEGG and NR annotation. The accumulations of some differentially expressed miRNAs and their corresponding target genes were determined in developing common buckwheat seed by qRT-PCR assays. In addition, the cleavage sites between 4 differentially expressed miRNAs and their target TFs were confirmed by 5′ RLM-RACE. As the first study to identify miRNA in common buckwheat, it not only provided insights into the miRNA participated in common buckwheat seed development, but also laid the foundation for further miRNA study in common buckwheat.

Author Contributions

Hongyou Li designed and participated in the study, analyzed the data and wrote the manuscript. Xiaoqian Sun, Chao Ma, and Hengling Meng participated in experiments. Qingfu Chen revised the manuscript.

References

Allen E, Z Xie, AM Gustafson, JC Carrington (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121:207–221

Bai B, B Shi, N Hou, Y Cao, Y Meng, H Bian, M Zhu, N Han (2017). microRNAs participate in gene expression regulation and phytohormone cross-talk in barley embryo during seed development and germination. *BMC Plant Biol* 17; Article 150

Bulgakov VP, TV Avramenko (2015). New opportunities for the regulation of secondary metabolism in plants: Focus on microRNAs. *Biotechnol Lett* 37:1719–1727

Cui H, MP Levesque, T Vernoux, JW Jung, AJ Paquette, KL Gallagher, JN Wang, I Blilou, B Scheres, PN Benfey (2007). An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants. *Science* 316:421–425

D'Ario M, S Griffiths-Jones, M Kim (2017). Small RNAs: Big impact on plant development. *Trends Plant Sci* 22:1056–1068

DeBoer K, S Melser, J Sperschneider, LG Kamphuis, G Garg, LL Gao, K Frick, KB Singh (2019). Identification and profiling of narrow-leafed lupin (*Lupinus angustifolius*) microRNAs during seed development. *BMC Genomics* 20; Article 135

Duan P, S Ni, J Wang, B Zhang, R Xu, Y Wang, H Chen, X Zhu, Y Li (2015). Regulation of *OsGRF4* by *OsmiR396* controls grain size and yield in rice. *Nat Plants* 2; Article 15203

Gao J, T Wang, M Liu, J Liu, Z Zhang (2017). Transcriptome analysis of filling stage seeds among three buckwheat species with emphasis on rutin accumulation. *PLoS One* 12; e0189672

Gao LC, YF Wang, ZY Zhu, H Chen, RH Sun, YS Zheng, DD Li (2019). EgmiR5179 from the mesocarp of oil palm (*Elaeis guineensis Jacq.*) regulates oil accumulation by targeting NAD transporter 1. *Ind Crop Prod* 137:126–136

Hu R, Z Li, S Xiang, Y Li, P Yi, M Xiao, X Zhang (2019). Comparative MicroRNA profiling reveals the cold response mechanisms in two contrasting tobacco cultivars. *Intl J Agric Biol* 22:757–762

Huang J, J Deng, T Shi, Q Chen, C Liang, Z Meng, L Zhu, Y Wang, F Zhao, S Yu, Q Chen (2017). Global transcriptome analysis and identification of genes involved in nutrients accumulation during seed development of rice tartary buckwheat (*Fagopyrum tararicum*). *Sci Rep* 7; Article 11792

Huang J, Z Li, D Zhao (2016). Deregulation of the *OsmiR160* target gene *OsARF18* causes growth and developmental defects with an alteration of auxin signaling in rice. *Sci Rep* 6; Article 29938

Jiao YQ, YH Wang, DW Xue, J Wang, MX Yan, GF Liu, GJ Dong, DL Zeng, ZF Lu, XD Zhu, Q Qian, JY Li (2010). Regulation of *OsSPL14* by OsmiR156 defines ideal plant architecture in rice. *Nat Genet* 42:541–544

Kang Z, CZ Zhang, JL Zhang, P Jin, J Zhang, GC Du, J Chen (2014). Small RNA regulators in bacteria: Powerful tools for metabolic engineering and synthetic biology. *Appl Microbiol Biotechnol* 98:3413–3424

Langmead B, C Trapnell, M Pop, SL Salzberg (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10; Article R25

Li C, BH Zhang (2016). MicroRNAs in control of plant development. *J Cell Physiol* 231:303–313

Li D, Z Liu, L Gao, L Wang, M Gao, Z Jiao, H Qiao, J Yang, M Chen, L Yao, R Liu, Y Kan (2016a). Genome-wide identification and characterization of microRNAs in developing grains of *Zea mays* L. *PLoS One* 11; Article e0153168

Li SC, FY Gao, KL Xie, XH Zeng, Y Cao, J Zeng, ZS He, Y Ren, WB Li, QM Deng, SQ Wang, AP Zheng, J Zhu, HN Liu, LX Wang, P Li (2016b). The OsmiR396c-*OsGRF4-OsGIF1* regulatory module determines grain size and yield in rice. *Plant Biotechnol J* 14:2134–2146

Li N, R Xu, Y Li (2019a). Molecular networks of seed size control plants. *Annu Rev Plant Biol* 70:435–463

Li H, Q Lv, J Deng, J Huang, F Cai, C Liang, Q Chen, Y Wang, L Zhu, X Zhang, Q Chen (2019b). Transcriptome analysis reveals key seed-development genes in common buckwheat (*Fagopyrum esculentum*). *Intl J Mol Sci* 20; Article 4303

Li T, L Ma, YK Geng, CY Hao, X Chen, X Zhang (2015). Small RNA and degradome sequencing reveal complex roles of miRNAs and their targets in developing wheat grains. *PLoS One* 10; Article e0139658

Luo M, ES Dennis, F Berger, WJ Peacock, A Chaudhury (2005). *MINISEED3* (*MINI3*) a WRKY family gene and *HAIKU2* (*IKU2*) a leucine-rich repeat (LRR) *KINASE* gene are regulators of seed size in *Arabidopsis*. *Proc Natl Acad Sci USA* 102:17531–17536

Ma C, J Zhang, J Yuan, J Guo, Y Xiong, Y Feng (2019). Differential expression of microRNAs is responsive to drought stress and exogenous methyl jasmonate in wheat (*Triticum aestivum*). *Intl J Agric Biol* 22:475–486

Ma X, X Zhang, K Zhao, F Li, K Li, L Ning, J He, Z Xin, D Yin (2018). Small RNA and degradome deep sequencing reveals the roles of microRNAs in seed expansion in peanut (*Arachis hypogaea* L.). *Front Plant Sci* 9; Article 349

Na GN, XP Mu, P Grabowski, J Schmutz, CF Lu (2019). Enhancing microRNA167A expression in seed decreases the α-linolenic acid content and increases seed size in *Camelina sativa*. *Plant J* 98:346–358

Pan JW, DH Huang, ZL Guo, Z Kuang, H Zhang, XY Xie, ZF Ma, SP Gao, MT Lerdau, CC Chu, L Li (2018). Overexpression of microRNA408 enhances photosynthesis growth and seed yield in diverse plants. *J Integr Plant Biol* 60:323–340

Peng T, M Qiao, H Liu, S Teotia, Z Zhang, Y Zhao, B Wang, D Zhao, L Shi, C Zhang, B Le, K Rogers, C Gunasekara, H Duan, Y Gu, L Tian, J Nie, J Qi, F Meng, L Huang, Q Chen, Z Wang, J Tang, X Tang, T Lan, X Chen, H Wei, Q Zhao, G Tang (2018). A resource for inactivation of microRNAs using short tandem target mimic technology in model and crop plants. *Mol Plant* 11:1400–1417

Peng T, H Sun, M Qiao, Y Zhao, Y Du, J Zhang, J Li, G Tang, Q Zhao (2014). Differentially expressed microRNA cohorts in seed development may contribute to poor grain filling of inferior spikelets in rice. *BMC Plant Biol* 14; Article 196

Romualdi C, S Bortoluzzi, F D'Alessi, GA Danieli (2003). IDEG6: A web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol Genomics* 12:159–162

Samad AFA, M Sajad, N Nazaruddin, A Izzat, IA Fauzi, AMA Murad, Z Zainal, I Ismanizan (2017). MicroRNA and transcription factor: Key players in plant regulatory network. *Front Plant Sci* 8; Article 565

Savadi S (2017). Molecular regulation of seed development and strategies for engineering seed size in crop plants. *Plant Growth Regul* 84:401–422

Shi TX, RY Li, QJ Chen, Y Li, F Pan, QF Chen (2017). *De novo* sequencing of seed transcriptome and development of genic-SSR markers in common buckwheat (*Fagopyrum esculentum*). *Mol Breed* 37; Article 147

Silva EMD, GFFE Silva, DB Bidoia, MD Silva Azevedo, FAD Jesus, LE Pino, LEP Peres, E Carrera, I López-Díaz, FTS Nogueira (2017). microRNA159-targeted *SlGAMYB* transcription factors are required for fruit set in tomato. *Plant J* 92:95–109

Sun LJ, XJ Li, YC Fu, ZF Zhu, LB Tan, FX Liu, XY Sun, XW Sun, CQ Sun (2013). *GS6* a member of the GRAS gene family negatively regulates grain size in rice. *J Integr Plant Biol* 55:938–949

Tian Y, Y Tian, X Luo, T Zhou, Z Huang, Y Liu, Y Qiu, B Hou, D Sun, H Deng, S Qian, K Yao (2014). Identification and characterization of microRNAs related to salt stress in broccoli, using high-throughput sequencing and bioinformatics analysis. *BMC Plant Biol* 14; Article 226

Wang J, H Jian, T Wang, L Wei, J Li, C Li, L Liu (2016). Identification of microRNAs actively involved in fatty acid biosynthesis in developing *Brassica napus* seeds using high-throughput sequencing. *Front Plant Sci* 7; Article 1570

Wang YM, Y Ding, DW Yu, W Xue, JY Liu (2015). High-throughput sequencing-based genome-wide identification of microRNAs expressed in developing cotton seeds. *Sci Chin Life Sci* 58:778–786

Wei J, XJ Lu, XL Zhang, M Mei, XL Huang (2017). Functions of microRNA in seed development dormancy and germination processes. *Hereditas* 39:14–21

Wei W, G Li, X Jiang, Y Wang, Z Ma, Z Niu, Z Wang, X Geng (2018). Small RNA and degradome profiling involved in seed development and oil synthesis of *Brassica napus*. *PLoS One* 13; Article e0204998

Xu Y, S Zhu, F Liu, W Wang, X Wang, G Han, B Cheng (2018). Identification of *Arbuscular Mycorrhiza* fungi responsive microRNAs and their regulatory network in maize. *Intl J Mol Sci* 19; Article 3201

Yasui Y, H Hirakawa, M Ueno, K Matsui, T Katsube-Tanaka, SJ Yang, J Aii, S Sato, M Mori (2016). Assembly of the draft genome of buckwheat and its applications in identifying agronomically useful genes. *DNA Res* 23:215–224

Yu L, R Guo, Y Jiang, X Ye, Z Yang, Y Meng, C Shao (2019). Genome-wide identification and characterization of novel microRNAs in seed development of soybean. *Biosci Biotechnol Biochem* 83:233–242

Zhang B, L Liu, H Wang, T Guan, Y Huang, C Liu (2019). Identification of *Plasmopara viticola* responsive microRNAs in grapevine by deep-sequencing. *Intl J Agric Biol* 22:801–807

Zhang JP, Y Yu, YZ Feng, YF Zhou, F Zhang, YW Yang, MQ Lei, YC Zhang, YQ Chen (2017a). MiR408 regulates grain yield and photosynthesis via a phytocyanin protein. *Plant Physiol* 175:1175–1185

Zhang H, J Zhang, J Yan, F Gou, Y Mao, G Tang, JR Botella, JK Zhu (2017b). Short tandem target mimic rice lines uncover functions of miRNAs in regulating important agronomic traits. *Proc Natl Acad Sci USA* 114:5277–5282

Zhang YC, Y Yu, CY Wang, ZY Li, Q Liu, J Xu, JY Liao, XJ Wang, LH Qu, F Chen, P Xin, C Yan, J Chu, HQ Li, YQ Chen (2013). Overexpression of microRNA OsmiR397 improves rice yield by increasing grain size and promoting panicle branching. *Nat Biotechnol* 31:848–852

Zhang Z, L Jiang, J Wang, P Gu, M Chen (2015). MTide: An integrated tool for the identification of miRNA-target interaction in plants. *Bioinformatics* 31:290–291

Zhao YF, T Peng, HZ Sun, S Teotia, HL Wen, YX Du, J Zhang, JZ Li, GL Tang, HW Xue, QZ Zhao (2019). miR1432-*OsACOT* (Acyl-CoA thioesterase) module determines grain yield *via* enhancing grain filling rate in rice. *Plant Biotechnol J* 17:712–723