Exploration of miRNA-mediated fertility regulation network of cytoplasmic male sterility during flower bud development in soybean

Xianlong Ding^{1,2} · Hao Zhang¹ · Hui Ruan¹ · Yanwei Li¹ · Linfeng Chen¹ · Tanliu Wang¹ · Ling Jin¹ · Xiaoqiang Li¹ · Shouping Yang¹ · Junyi Gai¹

Received: 10 September 2018 / Accepted: 16 December 2018 / Published online: 2 January 2019
© King Abdulaziz City for Science and Technology 2019

Abstract

Cytoplasmic male sterility (CMS) plays an important role in the production of soybean hybrid seeds. MicroRNAs (miRNAs) are a class of non-coding endogenous ~21 nt small RNAs that play crucial roles in flower and pollen development by targeting genes in plants. To dissect the function of miRNAs in soybean CMS, a total of 558 known miRNAs, 10 novel miRNAs, and 466 target genes were identified in flower buds of the soybean CMS line NJCMS1A and its restorer line NJCMS1C through small RNA sequencing and degradome analysis. In addition, miRNA-mediated editing events were also observed, and the two most frequently observed editing types (A → G and C → U) were validated by cloning and sequencing. And as the base editing occurred, some targets were filtered, such as gma-miR2118b-P6GT with *Glyma.08G122000.2*. Further integrated analysis of transcriptome and small RNA found some miRNAs and their targets' expression patterns showing a negative correlation, such as gma-miR156b/*GmSPL9a* and gma-miR4413b/*GmPPR*. Furthermore, opposite expression pattern was observed between gma-miR156b and *GmSPL9* during early stage of flower bud development. Taken together, the regulatory network of gma-miR156b/*GmSPL9* and gma-miR4413b/*GmPPR* with flower bud development in soybean CMS was developed. Most importantly, previous reports showed that these targets might be related to pollen development and male sterility, suggesting that both conserved and species-specific miRNAs might act as regulators of flower bud development in soybean CMS. These findings may provide a better understanding of the miRNA-mediated regulatory networks of CMS mechanisms in soybean.

Keywords Soybean (*Glycine max* (L.) Merr.) · Cytoplasmic male sterility · Fertility regulation network · miRNA · Target gene

Abbreviations

CMS	Cytoplasmic male sterility
MFE	Minimum free energies
MFEI	Minimal folding energy indices
miRNA	MicroRNA
PPR	Pentatricopeptide repeat-containing
qRT-PCR	Quantitative real time PCR

Rf	Restorer of fertility
SPL	Squamosa promoter-binding protein-like

Introduction

Soybean is one of the most important oil crops in the world. However, soybean yields are relatively low and the utilization of heterosis is one of the effective ways to improve them. Cytoplasmic male sterility (CMS) is a simple and efficient pollination control system that plays an important role in the production of hybrid seed. CMS is caused by mitochondrial genes with coupled nuclear genes, and CMS-based hybrid seed technology uses a three-line system: the CMS line, the maintainer line, and the restorer line (Chen and Liu 2014). The CMS line is propagated by crossing with the maintainer line and by crossing the CMS line with the restorer line, the

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1543-1>) contains supplementary material, which is available to authorized users.

✉ Shouping Yang
spyung@126.com

✉ Junyi Gai
sri@njau.edu.cn

Extended author information available on the last page of the article

male fertility of the F_1 plant can be restored by the *Rf* (restorer of fertility) gene(s) that come from the nuclear genome of the restorer line. Whether it is the production of CMS or F_1 fertility restoration, a series of changes will occur at both the molecular and physiological levels. In this process, the interaction between genes is very obvious. Therefore, the molecular mechanism of CMS is a hot research topic and there may be a high degree of epigenetic regulation involvement.

MicroRNAs (miRNAs) are a class of non-coding endogenous ~21 nt small RNAs (Bartel 2004) that play vital regulatory roles in plant growth and development, such as flowering-time regulation (Nie et al. 2015), leaf development (Samad et al. 2018), accumulation of anthocyanins (Liu et al. 2017; Zhao et al. 2017), response to the phytohormone abscisic acid (Duan et al. 2016), root development (Liu et al. 2018), cotton fiber elongation (Wang et al. 2016a), pollen or flower bud development (Shen et al. 2011; Yang et al. 2013; Jiang et al. 2014; Ding et al. 2016; Zhang et al. 2016), and so on. Among all of the miRNA families, MIR156 is one of the most conserved families in plants (Wang et al. 2016b) and regulates the vegetative phase change floral transition, integrated into flowering pathways by target squamosa promoter binding protein-like (*SPL*) transcription factors (Spanudakis and Jackson 2014). Moreover, the miR156 targeted or non-targeted *SPL* genes function together to redundantly secure male fertility in *Arabidopsis* (Xing et al. 2010). Compared with conserved miRNA, the expression of species-specific miRNA is relatively low, and they can be induced to function under certain special conditions (Tang 2010).

Our previous results have shown that epigenetic modification, such as DNA methylation and circular RNAs, are involved in flower bud development of soybean CMS (Li et al. 2017b; Chen et al. 2018). miRNAs, as a form of epigenetic modification, were purported to take part in the process of male sterility in plant (Xing et al. 2010; Yu et al. 2013). In this study, a number of miRNAs and their targets in soybean were identified by comparing a CMS line (NJCMS1A) and its restorer line (NJCMS1C), based on small RNA sequencing and degradome analysis. Most importantly, the regulatory network of gma-miR156b/*GmSPL9* and gma-miR4413b/*GmPPR* with flower bud development in soybean CMS was found by small RNA sequencing, transcriptome and qRT-PCR comparative analysis. These results might shed light on the potential regulatory role of miRNAs in controlling flower bud development of soybean CMS.

Materials and methods

Plant materials, sample collection, and total RNA extraction

The soybean cytoplasmic male-sterile line NJCMS1A was developed through consecutive backcross procedures with the cultivar N8855 as donor parent and N2899 (hereafter designated as NJCMS1B) as recurrent parent (Gai et al. 1995; Ding et al. 1999, 2002). “Zhongdou 5” (hereafter designated as NJCMS1C for further evaluation of the genetic mechanism of male fertility restoration) was identified as the restorer of NJCMS1A in a previous study (Yang et al. 2007). NJCMS1A and NJCMS1C were planted in the summer of 2013 at Jiangpu Experimental Station, National Center for Soybean Improvement, Nanjing Agricultural University, Nanjing, Jiangsu, China, for sample collection for small RNA sequencing and degradome sequencing. The male-sterile plants and fertile plants were identified through three methods, namely, the dehiscence of anthers, germination rate of pollen, and performance of plants at maturity. Cytological observation showed that the male abortion of NJCMS1A occurred mainly at the early binucleate pollen stage (Fan 2003). Because it is very difficult to judge the precise developmental stage of pollen from the appearance of the flower buds in soybean, during the flowering period in the summer of 2013, flower buds of different sizes up to and including the abortion stage were collected and mixed from NJCMS1A and NJCMS1C plants, and then immediately frozen in liquid nitrogen and stored at -80°C for further use. Flower buds were collected from NJCMS1A and NJCMS1C in the summer of 2015 for quantitative real-time PCR (qRT-PCR). Total RNAs from the flower buds of NJCMS1A and NJCMS1C were extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s protocol.

Small RNA sequencing library construction and bioinformatics analysis

Two small RNA libraries were constructed and sequenced by the Beijing Genomics Institute (BGI, Shenzhen, Guangdong Province, China) using the Illumina HiSeq 2000 System (TruSeq SBS KIT-HS V3, Illumina, Santa Clara, CA, USA). After sequencing, raw sequencing reads were processed into clean reads by filtering out adaptor contaminants, oversized insertions, low-quality reads, poly(A) tags, and small tags (< 18 nt). The small RNA tags were mapped to the soybean genome (Wm82.a2.v1) using SOAP to analyze their expression and distribution on the genome. All small RNA sequences were used to

query the RNA sequences for deposited repeats (RepeatMasker output) in GenBank (NCBI GenBank). Rfam (V 11.0) was used to separate the small RNAs that matched rRNA, scRNA, snoRNA, snRNA, and tRNA and remove the degraded mRNA fragments (Wm82.a2.v1) in the small RNA tags that aligned to exons and introns. Next, to identify conserved miRNAs, clean reads were compared with known miRNAs of soybean deposited at miRBase 21.0 (<http://www.mirbase.org/>). Finally, the other sequences that did not map to known miRNAs and other kinds of small RNA were referred to as unannotated sequences for identification of new members of known miRNA families and for novel miRNA prediction.

Potentially novel miRNAs were identified using MIREAP (<https://source-forge.net/projects/mireap/>), and their secondary structures were predicted by the mfold web server (<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>) (Zuker 2003). The criteria used for selecting novel miRNAs must meet the following characteristics based on Meyers et al. (2008), Zhang et al. (2006), Kozomara and Griffiths-Jones (2014), and miRBase 21.0: (1) the candidate miRNA-5p and miRNA-3p are derived from opposite stem arms with minimal matched nucleotide pairs exceeding 16 nt and with maximal size differences of up to 4 nt; (2) the most abundant reads from each arm of the precursor must pair in the mature miRNA duplex with a 2-nt 3' overhang; (3) the number of asymmetric bulges within the miRNA-5p/miRNA-3p duplex must be one or fewer, and the size of the asymmetric bulges must be two bases or smaller; (4) the miRNA-5p and miRNA-3p are simultaneously on the two arms of pre-miRNA secondary structures, and one of the mature miRNAs must be no fewer than five; (5) the candidate miRNA precursor must have high negative minimum free energy (MFE) and minimum folding energy index (MFEI), with $MFE < -0.2$ kcal/mol/nt and $MFEI > 0.85$.

Degradome library construction, data analysis, and target identification

Two degradome libraries were constructed according to the methods of Addo-Quaye et al. (2008) and German et al. (2008) and sequenced on the Illumina HiSeq 2000 (TruSeq SBS KIT-HS V3, Illumina, BGI, Shenzhen, Guangdong Province, China). Alignments with scores up to 4.5, where G:U pairs scored 0.5 and no mismatches were found at the site between the 10th and 11th nucleotides of the corresponding miRNAs, were considered potential targets.

Quantitative real-time PCR validation

Stem-loop qRT-PCR and qRT-PCR were carried out to validate differential expression levels of miRNAs and

miRNA targets, respectively (Chen et al. 2005). All primers (Table S1) were designed based on the mature miRNA and mRNA sequences and synthesized commercially (Invitrogen, Shanghai, China). According to the procedures provided in the iScript Select cDNA Synthesis Kit (containing GSP enhancer solution, BIO-RAD, USA), 1 µg of total RNA was reverse transcribed by iScript reverse transcriptase using stem-loop primers. The qRT-PCR analysis was carried out using iTaq Universal SYBR Green Supermix (BIO-RAD, USA) on a Bio-Rad CFX96 instrument (CFX96 Touch, BIO-RAD, USA). All reactions were run with three biological replicates, and gma-miR1520d (Kulcheski et al. 2010) was used as the internal control gene. The relative expression levels of miRNAs and genes were quantified using the $2^{-\Delta\Delta C_t}$ method. Student's *t* test was performed to compare miRNA or mRNA differences in expression between NJCMS1A and NJCMS1C. Expression levels that significantly differed ($p < 0.05$) according to Student's *t* test were labeled as: "*", "**", and "***", which represent $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively, which indicated significant difference levels between NJCMS1A and NJCMS1C.

Results

Overview of small RNA library data sets

Two small RNA libraries were constructed using total RNA obtained from flower buds of the CMS line NJCMS1A and its restorer line NJCMS1C and sequenced by Illumina Solexa high-throughput sequencing technology. After removing adaptor contaminants and low-quality reads etc., a total of 11,974,839 and 12,072,909 clean reads were obtained from the two libraries, respectively, ranging from 18- to 30-nt in length (Fig. 1a). Using the SOAP software, more than 80% small RNA tags were perfectly mapped to the soybean genome (Table 1).

Upon analyzing the small RNA tags between NJCMS1A and NJCMS1C, we found that NJCMS1A and NJCMS1C comprised 14.86% and 18.31% of the total sRNAs (Fig. 2a), respectively. Only 12.64% of unique sRNAs were shared between the CMS line and its restorer line (Fig. 2b). Moreover, the NJCMS1C-specific unique sRNAs were greater than those of NJCMS1A, which means that most of the unique sRNAs found in NJCMS1C flower buds were different from those in NJCMS1A (Fig. 2b). The small RNAs annotation for the two samples is shown in Fig. 2c, d. The majority of small RNAs were 21–24 nt, and 24-nt small RNAs were the most abundant (Fig. 1a), which is in accordance with the previously reported miRNAs of soybean (Song et al. 2011).

Fig. 1 Length distribution of small RNA sequences and degradome-seq clean tags in flower buds of NJCMS1A and NJCMS1C. **a** Length distribution of small RNA sequences derived from the small RNA

libraries. **b** Length distribution of degradome-seq clean tags derived from the degradome libraries

Table 1 Summary of small RNA sequences and degradome sequences

Type	Small RNA-seq		Degradome-seq	
	NJCMS1A	NJCMS1C	NJCMS1A	NJCMS1C
Statistical information				
Total reads	12,190,078 (100.00%)	12,267,565 (100.00%)	12,351,703 (100.00%)	12,669,765 (100.00%)
Clean reads	11,974,839 (98.93%)	12,072,909 (99.15%)	12,288,476 (99.49%)	12,613,736 (99.55%)
Genome mapping statistics				
Total count of clean tags				
Total tags	11,974,839 (100.00%)	12,072,909 (100.00%)	12,288,476 (100.00%)	12,617,159 (100.00%)
Mapping to genome	9,808,568 (81.91%)	9,985,396 (82.71%)	10,926,341 (88.92%)	11,244,308 (89.12%)
Unique count of clean tags				
Total tags	3,668,753 (100.00%)	4,339,928 (100.00%)	6,332,697 (100.00%)	6,482,757 (100.00%)
Mapping to genome	2,682,237 (73.11%)	3,204,022 (73.83%)	5,494,191 (86.76%)	5,638,108 (86.97%)

Identification of known miRNAs

The mappable small RNA sequences were aligned with the known soybean miRNAs using miRBase 21.0. A total of 499 known pre-miRNAs corresponding to 558 mature miRNAs that belong to 222 families were detected (Table S2, S3). The sequences of some miRNAs were tagged with “D” (Table S2), representing that these were either shorter or longer than the mature miRNAs in miRBase 21.0; these can be classified as isomiRNAs and may share the same target genes with the mature miRNAs.

Identification of novel miRNAs

To identify novel miRNAs, all unannotated and mappable small RNAs were aligned to known plant miRNAs in miRBase 21.0. The small RNAs that could not be mapped

to known plant miRNAs were classified as novel miRNAs. The five criteria described in “[Materials and methods](#)” were utilized to increase predictive accuracy. As a result, ten novel miRNAs (five pairs) that belong to four families were identified in this study (Table 2 and Table S3, S4). Precursors of these novel miRNAs were identified by MIREAP and varied from 85 to 93 nt in length, with MFE values ranging from -0.47 to -0.57 kcal/mol/nt. The MFEI values ranged from 1.15 to 1.65 (Table S4), all of them being greater than 0.97 and the average being 1.36. The length of all of the mature miRNAs was 21 nt. Some of the novel miRNAs, such as novel_mir_003a to novel_mir_003b, shared the same mature sequence but had different precursors that came from different loci. The secondary structure for the precursor of novel_mir_003a was selected as an example to be shown in Fig. S1.

Fig. 2 Distribution of total sRNAs (tags) and unique sRNAs (tags) among small RNA libraries and degradome libraries. **a, b** Distribution of total sRNAs and unique sRNAs in small RNA libraries of

NJCMS1A and NJCMS1C. **c, d** Annotation of sRNAs in small RNA libraries of NJCMS1A and NJCMS1C. **e, f** Distribution of total tags and unique tags in degradome libraries of NJCMS1A and NJCMS1C

Table 2 Novel miRNAs identified in NJCMS1A and NJCMS1C

miRNA Name	Novel miRNA Sequence	Length (nt)	NJC-MS1A Count	NJC-MS1C Count	3p/5p ^a
novel_mir_001-5p	TGGAGAAGCAGGGCACATGCT	21	78	136	5p
novel_mir_001-3p	CTTGTGTCCTACTTCTCCAGC	21	4	11	3p
novel_mir_002-5p	GAGCAATTCTCCTTTGGCAGA	21	2	0	5p
novel_mir_002-3p	TGCCAAAGGAGAATTGTCCTG	21	6	0	3p
novel_mir_003a-5p	TTTATCGGTAACATCATCATC	21	261	0	5p
novel_mir_003a-3p	TGATGATGCTACTGATAAATA	21	4	0	3p
novel_mir_003b-5p	TTTATCGGTAACATCATCATC	21	261	0	5p
novel_mir_003b-3p	TGATGATGCTACTGATAAATA	21	4	0	3p
novel_mir_004-5p	TAAACGTTTGATCCCTTGTAT	21	8	0	5p
novel_mir_004-3p	ACAAGGGATCAAACATTTAAG	21	1	0	3p

^a Arm of this mature sequence on the predicted secondary structure

Overview of miRNA base edits

In the present study, miRNAs which might have base edit were detected by aligning unannotated sRNA tags with mature miRNAs from miRBase21.0. A total of 174 miRNA members exhibited base edit and the miRNA editing frequency was 0.13–100.00% (Table S5). Some miRNAs, such as gma-miR1516a-3p, were not found by sequencing, but

its editing event was observed in this study (Table S5-1). The miRNA editing events had truly occurred in soybean, as confirmed by our sequencing results (Fig. S2). Most of the miRNA editing events occurred at nucleotide positions of 5–17 and the miRNA editing patterns were similar between the two lines, but the editing frequency at position of 11 is slightly lower (Fig. 3a). The most dominant nucleotide substitution types were A → G, C → U and G → U (Fig. 3b).

Fig. 3 Summary and validation of miRNA editing events in flower buds of NJCMS1A and NJCMS1C. **a** Summary of miRNA editing frequency at each miRNA nucleotide position in both NJCMS1A and NJCMS1C. **b** Summary of miRNA editing types and their frequency in both NJCMS1A and NJCMS1C. **c, d** Validation of miRNA editing types inferred from miRNA precursors and mature miRNAs by high-throughput sequencing and cloning. Sequencing chromatograms from

the two observed editing types (C→T, A→G) of gma-miR159e and gma-miR4403 are shown. The edited positions are highlighted in pink. The upper panel indicates the part of precursor miRNAs cloned from genomic DNA, and the lower panel indicates the mature miRNAs cloned from cDNA reverse transcribed with stem-loop RT primers. Primers are listed in Table S1

To validate the occurrence of the miRNA editing, the two most frequently observed edited types ($A \rightarrow G$ and $C \rightarrow U$) were examined. The precursor miRNA sequence was cloned from the soybean genome, and the mature miRNA sequence was cloned from the cDNA reverse transcribed by the stem-loop reverse primers. Both sequencing and validate results confirmed that nucleotide substitutions had truly occurred during the mature miRNA processing (Table S5 and Fig. 3c, d). However, the reason for the variant editing type and their function in mature miRNA processing or gene expression regulation remain elusive and warrant further study.

Comparative analysis of miRNAs

Throughout all the identified miRNAs in flower buds of NJCMS1A and NJCMS1C, 76 differentially expressed miRNAs (including 74 known miRNAs and two novel miRNAs) with more than two-fold relative change were identified (Table S6). Among the differentially expressed miRNAs, 28 miRNAs were upregulated in NJCMS1A, and the remaining miRNAs showed higher expression levels or were only expressed in flower bud of NJCMS1C (Table S6). For example, gma-miR4393a and gma-miR5667 were only expressed in NJCMS1A, and gma-miR4346, gma-miR4372a, gma-miR4400, gma-miR4994-5p, and gma-miR5785 were only detected in NJCMS1C. In particular, gma-miR4346 showed a high relative expression level and may have a unique function in the process of pollen development in NJCMS1C. Of the differentially expressed novel miRNAs, most of them were only expressed in NJCMS1A or NJCMS1C (Table S6).

To examine miRNA expression, ten selected differentially expressed miRNAs were assayed by stem-loop qRT-PCR. As showed in Fig. 4, nine of these were consistent with the sequencing reads. However, gma-miR6300 was found to be upregulated in NJCMS1C, but was enriched in NJCMS1A based on the high-throughput sequencing analysis and this result warrants further study.

Degradome library construction and comparative analysis of miRNA targets between NJCMS1A and NJCMS1C

Here, we performed degradome sequencing and analysis to search for the miRNA targets in NJCMS1A and NJCMS1C. A total of 12,288,476 and 12,617,159 clean reads were generated from the NJCMS1A and NJCMS1C libraries, respectively (Table 1). More than 98.00% of the tags were 20 or 21 nt in length (Fig. 1b), as expected from the normal length distribution peak of degradome fragments. Of the unique tags, 41.89% and 43.24% were from NJCMS1A and NJCMS1C, respectively, and only 14.87% of them were shared between NJCMS1A and NJCMS1C (Fig. 2e, f).

Fig. 4 Verification results of differentially expressed miRNAs between NJCMS1A and NJCMS1C. The y axis indicates the miRNA relative expression level (log2) generated from qRT-PCR analysis and high-throughput sequencing. The results were obtained from three biological replicates, and the error bars indicate the standard error of the mean of $2^{-\Delta\Delta C_t}$, with NJCMS1C as a control

To better understand the function of miRNA in flower bud development, we performed a conjoint analysis of the miRNA and degradome together. A total of 466 distinct transcripts targeted by 200 miRNAs were detected (Table S7 and Fig. S3). These transcripts included *SPL* family protein, *GAMYB* protein, auxin response factor, homeobox-leucine zipper protein, nuclear transcription factor Y, *GRAS* family transcription factor, and transcription factor *APETALA2*, which have essential roles in gene regulation (Table S7). Most importantly, a set of pentatricopeptide-repeat-containing (PPR) proteins were targeted by gma-miR1508, gma-miR160, gma-miR171o, gma-miR4413b and gma-miR5674, which may be a group of *Rf* proteins that suppress CMS gene expression through posttranscriptional mechanisms such as editing, cleavage, and degradation of the target mRNAs (Table S8; Chen and Liu 2014).

To determine if the base edit in positions 2–8 and 10–11 of a mature miRNA can change its target recognition, a combined analysis of the degradome library and base-edited miRNAs was conducted. As a result, 122 distinct transcripts targeted by 307 base-edited miRNAs were detected (Table S9). Most of the targets of the base-edited miRNAs were the same as for the corresponding miRNAs for which no base editing had occurred (Table S10). However, some targets were filtered out by the base editing. For example, there were 75 targets of gma-miR156b and gma-miR156f, but as the base edited, only 61 targets were detected (Table S10). Interestingly, we also detected some new targets that were not targeted by miRNA when no base editing occurred, such as *Glyma.05G230700.2* to *Glyma.05G230700.4*, which were targeted by gma-miR166h-P8CT, gma-miR166k-P8CT, and

gma-miR166u-P8CT (Table S10), but not by gma-miR166h, gma-miR166k, and gma-miR166u. Furthermore, with base editing in position 10–11 of the mature miRNA, some miRNA targets were changed. For example, no targets were detected for gma-miR4403 in this study, but as a result of a base edit at position 11, one target (*Glyma.17G050600.1*) was found (Fig. S4 and Table S10).

Integrated analysis of miRNA and mRNA expression between NJCMS1A and NJCMS1C

To better understand the function of differentially expressed miRNAs in flower bud development of soybean, we performed an integrated analysis to identify inverse miRNA–mRNA expression based on small RNA and transcriptome comparative analysis (Li et al. 2017a). In consideration of the expression of some miRNAs and their targets were changed more significantly in sporogenic tissues but were submerged by the expression in other tissues of the flower buds, so a cutoff of > 1.5-fold change was used in this section as that of small RNA and transcriptome sequencing in Chinese cabbage CMS (Wei et al. 2015). As a result, 13 miRNAs (6 miRNA families) with their 7 inverse expressed target genes that had a fold change > 1.5 or < −1.5 were identified (Table S11 and Fig. S5). Among them, the conserved miRNA, gma-miR156b, was more than 1.5-fold greater in NJCMS1A. Twenty-three target genes were identified to be its targets in this study (Table S7). However, only one target showed a negative correlation with gma-miR156b (Table S11). Despite the conserved miRNAs, gma-miR4392 and gma-miR4413b were two soybean-specific miRNAs, which were also altered in NJCMS1A (Table S11). This result indicated species-specific miRNA may also participate in the flower bud development in plant. In addition, 3 conserved miRNA families, included gma-miR169, gma-miR1514 and gma-miR397, were more than 1.6-fold greater in the NJCMS1C. Of the targeted genes, 4 were found to be showing more than 2.0-fold higher in NJCMS1A (Table S11).

miRNA–target network underlying flower bud development of soybean CMS

Based on the integrated analysis of the expression profiles of miRNAs and their targets, six miRNA–targets pairs were found, in which miRNAs and target genes are negatively regulated at the expression level (Table S11 and Fig. S5). However, only half of the miRNA families were upregulated and suppressed the expression of three targets (Table S11 and Fig. S5). MIR156 is one of upregulated miRNA families and our present study found that 20 *GmSPL* genes were targeted by gma-miR156b (Table S7). Unfortunately, only *GmSPL9a* was downregulated in flower bud of soybean

CMS (Table S11 and Fig. 5). It is possible that the expression of other *GmSPL* genes was very sensitive to the stage of flower bud development. To prove this point, qRT-PCR analysis was conducted to detect the expression level of gma-miR156b and some of its targets included *GmSPL9b* and *GmSPL9a* (Table S12). Consistent with expectations, there was a negative relationship at expression level between gma-miR156b and *GmSPL9a* or *GmSPL9b* during early flower bud development (Fig. 6). In addition, our present study found a set of *GmPPR* genes was targeted by gma-miR4413b, which is a nonconserved miRNA and only identified in soybean now (miRBase 21.0 and Table S8). Most importantly, both transcriptome data and qRT-PCR analysis found gma-miR4413b was upregulated and one of the *GmPPR* genes was downregulated in soybean CMS, simultaneously (Fig. 5c, d). This finding indicated the possible existence of a crosstalk network between miRNA–target module and flower bud development of soybean CMS (Fig. 7).

Discussion

Exploration of miRNAs and their targets participating in flower bud development of soybean CMS

In plants, many miRNAs seem to be universally expressed across diverse tissues, including the flower bud or pollen. Previous reports have found 100 known miRNAs in the Rf3 and rf3 pollens of maize by solexa sequencing (Yu et al. 2013). In this study, 499 known pre-miRNAs corresponding to 558 mature miRNAs and 10 novel miRNAs were identified in flower buds of the CMS line and its restorer line in soybean. Among all these miRNAs, 76 differentially expressed miRNAs with more than two-fold change between the flower buds of NJCMS1A and NJCMS1C were identified (Table S6). We speculate that these miRNAs might be involved in a specific process of flower bud development. By degradome analysis, 466 targets were chosen and predicted to be cleaved by 200 miRNAs (Table S7). Among the identified targets of miRNAs, some have previously been shown to be involved in floral organ or pollen development, such as miR156 with *SPL* (Xing et al. 2010), miR172 with *APETALA2* (Aukerman and Sakai 2003), miR159 with *GAMYB* or *GAMYB*-like genes (Alonso-Peral et al. 2010), and so on.

Moreover, we detected 174 miRNA members in which base editing had occurred (Table S5). Wei et al. (2009) divided the miRNA sequence into three major areas; positions 2–8 and 9–16 of a mature miRNA were called the seed region and region A, respectively, and the remainder was denoted region B. miRNA editing events were first found

Fig. 5 Relative expression level of *gma-miR156b/GmSPL9a* and *gma-miR4413b/GmPPR* between NJCMS1A and NJCMS1C. **a, c, e** Relative expression level of *gma-miR156b/GmSPL9a* and *gma-miR4413b/GmPPR* in small RNA sequencing and transcriptome

analysis, respectively. **b, d, f** Relative expression level of *gma-miR156b/GmSPL9a* and *gma-miR4413b/GmPPR* in qRT-PCR analysis

in animals (Wei et al. 2009) and subsequently revealed to also occur during plant development (Yan et al. 2015). Previous report showed that the seed region is used for target recognition; if the seed region and region A or the seed region and region B work together, the miRNA can regulate different species-specific targets (Wei et al. 2009). As miRNAs function in plants by either target cleavage or translational repression, editing of the 10th/11th positions of miRNA will also lead to altered targets. To verify the above results, we combined the analysis of the degradome library and base-edited miRNAs (seed region and the 10th/11th positions of miRNA). The results showed that most of the targets of base-edited miRNAs were the same as those of the unedited miRNAs (Table S10). However, as the base editing occurred, some targets were filtered. For example, *Glyma.08G122000.2* was targeted by *gma-miR2118a-3p* and *gma-miR2118b-3p*, but when the base G at position 6 was changed to T, it no longer targeted

Glyma.08G122000.2 (Table S10). On the other hand, when the base C at position 11 of *gma-miR2118a-3p* and *gma-miR2118b-3p* was changed to T, we identified a new target (*Glyma.08G005700.1*) by degradome analysis (Table S10).

Several miRNA regulatory networks might be involved in regulation of flower bud development of soybean CMS

It is widely accepted that CMS is caused by mitochondrial genes with coupled nuclear genes, and the male fertility of the F_1 plants can be restored by the *Rf* gene(s) that comes from the nuclear genome of the restorer line (Chen and Liu 2014). In this process, the interaction between genes is very obvious, the nuclear genes might change their expression as target genes in response to the mitochondrial signaling, which subsequently affects flower bud development. In this study, altered expression of miRNAs and their targets genes

Fig. 6 Relative expression level of *gma-miR156b*, *GmSPL9a* and *GmSPL9b* in flower bud of three periods. **a** The flower buds of three periods that were sampled for qRT-PCR analysis. **b–d** Relative

expression level of *gma-miR156b*, *GmSPL9a* and *GmSPL9b* in flower bud of soybean CMS

was exhibited between NJCMS1A and NJCMS1C, demonstrating important roles of miRNAs regulation network in flower bud development of soybean CMS. For example, *GmNF-YA3/GmNF-YA5* and *GmLAC4* were target genes of *gma-miR169* and *gma-miR397* (Table S11), respectively. Transcriptome sequencing showed these target genes were

all more than twofold change expressed in NJCMS1A (Table S11), and previous studies reported their overexpression affects male gametogenesis and induces nondehiscent anthers in plant (Mu et al. 2013; Zhang et al. 2014; Nasrin et al. 2010). Based on the expression analysis of miRNAs and their corresponding targets, *gma-miR156b/GmSPL9* and *gma-miR4413b/GmPPR* regulatory network was selected as two candidate miRNAs–targets combinations that might be involved in flower bud development of soybean CMS.

MIR156 and *SPL* are one of the highly conserved miRNA and gene families in plant, respectively, and different members of MIR156 and *SPL* showed highly dynamic expression patterns at different stages of flower development in *Arabidopsis* (Xing et al. 2010). Former studies have shown that miR156 plays essential regulatory roles in floral development and male fertility by targeting *SPL* genes, and the loss of function of miR156/7-targeted *SPL* genes led to a semi-sterile phenotype in *Arabidopsis* (Xing et al. 2010). Yu et al. (2013) reported that miR156 had a higher expression level in the pollen of S-type CMS in maize. Moreover, it had higher expression levels in the anthers from the three anther developmental stages of the GMS mutant in cotton (Wei et al. 2013). In *Arabidopsis*, the pollen fertility of triple mutant *spl8 spl9 spl15*, *spl8 spl2 spl9*, and quadruple mutant *spl8 spl2 spl9 spl15* was lower than that of the *spl8* single mutant (Xing et al. 2010), which provided evidence that the *AtSPL9* do affect pollen fertility. Most importantly, overexpression of *AtSPL9* in the *spl8* mutant could restore its fertility partially (Xing et al. 2010), indicating that *AtSPL9* is crucial to pollen development. It has been reported that

Fig. 7 The speculated model of miRNA–targets regulatory network of soybean CMS during flower bud development. The up-regulated miRNAs and down-regulated mRNAs are in red and green, respectively. Solid lines indicate a pathway supported by experimental evidence in this study. Dotted lines show a speculated pathway supported by reports from literature. Pictures in left were morphology of anthers during flowering day in NJCMS1A and NJCMS1C, which exhibited anther indehiscence and dehiscence, respectively. Pictures in right were I_2 -KI staining of pollen during flowering day in NJCMS1A and NJCMS1C, whose pollen were sterile and fertile, respectively

SPL9 is essential for anther development, and overexpression of miR157 in cotton leads to microspore abortion and anther indehiscence under high-temperature stress by regulated *GhSPL9*, etc. (Ding et al. 2017). In this study, gma-miR156b negatively regulated *GmSPL9a* that displayed differential expression between NJCMS1A and NJCMS1C, indicating that miR156/*SPL9* regulatory network may also participate in flower bud development of soybean (Fig. 7).

In present study, a total of 4 *GmSPL9* genes were predicted to be targeted by gma-miR156b (Table S12). However, only *GmSPL9a* reached more than 1.5-fold expression difference between NJCMS1A and NJCMS1C (Table S11). Previous studies indicated that *SPL* gene plays a key role in early anther development by promoting other early anther genes required for cell division, specification, and differentiation, resulting in pollen mother cell and tapetum formation (Schiefthaler et al. 1999; Yang et al. 1999; Xing et al. 2010). So we also speculated that *GmSPL9* genes were also involved in flower bud development of soybean CMS in early stage of flower bud development. To verify this view, we used qRT-PCR to measure the expression of gma-miR156b, *GmSPL9a* and *GmSPL9b* at three periods during flower bud development of soybean. As a result, gma-miR156b was upregulated in the period I and II, and *GmSPL9a* and *GmSPL9b* were downregulated before period III and II (Fig. 6), respectively, which means that the targeted *GmSPL9* genes exhibit an opposite expression pattern with that of gma-miR156b during certain stage of flower bud development (Fig. 6). Thus, miR156 and *SPL9* may also be absolutely required for proper early flower bud development and cell proliferation and differentiation in soybean CMS, which warrants further study.

PPR proteins are classified as *Rf* proteins that suppress CMS gene expression through posttranscriptional mechanisms such as editing, cleavage, and degradation of the target mRNAs in many plants such as rice, *Brassica*, radish, and sorghum (Chen and Liu 2014; Brown et al. 2003; Hu et al. 2012; Iwabuchi et al. 1999; Jordan et al. 2010; Kazama and Toriyama 2003). Recently, Wang et al. (2016c) identified the *Rf* gene for soybean M-type CMS, *Rf-m*, which is located in a *PPR* gene-rich region on chromosome 16. This region contains 19 putative open reading frames (ORFs), and seven of them encode putative *PPR* proteins. Two of these (*Glyma.16G161900* and *Glyma.16G163100*) were targeted by gma-miR4413b. Gma-miR4413b is a soybean-specific miRNA and was upregulated (1.5-fold) in NJCMS1A, and its target genes were all *PPR* proteins (Table S8). Moreover, we found *Glyma.16G161900* was significantly downregulated in NJCMS1A according to the transcriptome data in our lab (Li et al. 2017a, b). Quantitative real-time PCR also showed that gma-miR4413b was upregulated in

NJCMS1A, and *Glyma.16G161900* was only expressed in NJCMS1C (Fig. 5d). In addition, the *Glyma.16G161900* was placed in the mitochondria (Table S8) and was homologous with *RPF1*, which is a *PPR* gene and supposed to play a role in fertility recovery of *Arabidopsis* CMS by editing *nad4* mRNA in mitochondria (Hölzle et al. 2011). Therefore, we speculated that the upregulation of gma-miR4413b reduced the expression level of *GmPPR* in soybean CMS, so that CMS gene can be expressed normally and lead to pollen sterility, eventually. All these results indicated that gma-miR4413b might play a specific function during the formation of soybean CMS (Fig. 7), which warrants further study.

Conclusion

In this study, a large number of miRNAs were identified during flower bud development in the soybean CMS line NJCMS1A and its restorer line NJCMS1C. By deep sequencing, 558 known miRNAs, 10 novel miRNAs, and 174 base-edited miRNA members were identified. Among the identified miRNAs, 76 differentially expressed miRNAs were discovered with greater than twofold changes between NJCMS1A and NJCMS1C. By degradome analysis, a total of 466 distinct transcripts targeted by 200 miRNAs and 122 distinct transcripts targeted by 307 base-edited miRNAs were detected. Small RNA sequencing, transcriptome and qRT-PCR comparative analysis found that gma-miR156b/*GmSPL9* and gma-miR4413b/*GmPPR* regulatory networks might be involved in flower bud development of soybean CMS. Further functional studies on these differentially expressed miRNAs will need to provide a better understanding of the miRNA-mediated regulation mechanisms during the CMS occurrence in the soybean N8855-CMS line.

Acknowledgements This work was supported by grants from the National Key R&D Program of China (2016YFD0101500, 2016YFD0101504), the National Hightech R&D Program of China (2011AA10A105), and the Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT_17R55, PCSIRT13073).

Author Contributions SPY, JYG and XLD conceived and designed the experiments. XLD, HZ, HR, YWL, LFC, TLW, LJ and XQL performed the experiments. XLD analyzed the data. XLD, HZ and HR contributed reagents/materials/analysis tools. XLD, YWL, LFC, TLW, LJ and XQL conceived the qRT-PCR experiments and analyzed the data. XLD, SPY conceived the experiments and wrote the manuscript. All authors read and approved the final manuscript.

Data Availability All of the small RNA-seq and degradome-seq data were submitted to the National Center for Biotechnology Information (NCBI) under the accession number PRJNA304685 and PRJNA504545.

Compliance with ethical standards

Conflict of interest All the authors declare that they do not have conflict of interest.

Ethical approval This article does not contain any study with human subjects or animals performed by any of the authors.

References

- Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ (2008) Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis* degradome. *Curr Biol* 18:758–762. <https://doi.org/10.1016/j.cub.2008.04.042>
- Alonso-Peral MM, Li JY, Li YJ, Allen RS, Schnippenkoetter W et al (2010) The microRNA159-regulated *GAMYB*-like genes inhibit growth and promote programmed cell death in *Arabidopsis*. *Plant Physiol* 154:757–771. <https://doi.org/10.1104/pp.110.160630>
- Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes. *Plant Cell* 15:2730–2741. <https://doi.org/10.1105/tpc.016238>
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–297. [https://doi.org/10.1016/S0092-8674\(04\)00045-5](https://doi.org/10.1016/S0092-8674(04)00045-5)
- Brown GG, Formanová N, Jin H, Wargachuk R, Dendy C et al (2003) The radish *Rfo* restorer gene of *Ogura cytoplasmic* male sterility encodes a protein with multiple pentatricopeptide repeats. *Plant J* 35:262–272. <https://doi.org/10.1046/j.1365-3113.2003.01799.x>
- Chen LT, Liu YG (2014) Male sterility and fertility restoration in crops. *Annu Rev Plant Biol* 65:579–606. <https://doi.org/10.1146/annurev-arplant-050213-040119>
- Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH et al (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33:e179. <https://doi.org/10.1093/nar/gni178>
- Chen LF, Ding XL, Zhang H, He TT, Li YW et al (2018) Comparative analysis of circular RNAs between soybean cytoplasmic male-sterile line NJCMS1A and its maintainer NJCMS1B by high-throughput sequencing. *BMC Genom* 19:663. <https://doi.org/10.1186/s12864-018-5054-6>
- Ding DR, Gai JY, Cui ZL, Yang SP, Qiu JX (1999) Development and verification of the cytoplasmic-nuclear male sterile soybean line NJCMS1A and its maintainer NJCMS1B. *Chin Sci Bull* 44:191–192. <https://doi.org/10.1007/BF02884752>
- Ding DR, Gai JY, Cui ZL, Qiu JX (2002) Development of a cytoplasmic-nuclear male-sterile line of soybean. *Euphytica* 124:85–91. <https://doi.org/10.1023/A:1015683526982>
- Ding XL, Li JJ, Zhang H, He TT, Han SH et al (2016) Identification of miRNAs and their targets by high-throughput sequencing and degradome analysis in cytoplasmic male-sterile line NJCMS1A and its maintainer NJCMS1B of soybean. *BMC Genom* 17:24. <https://doi.org/10.1186/s12864-015-2352-0>
- Ding YH, Ma YZ, Liu N, Xu J, Hu Q et al (2017) microRNAs involved in auxin signalling modulate male sterility under high-temperature stress in cotton (*Gossypium hirsutum*). *Plant J* 91: 977–994. <https://doi.org/10.1111/tj.13620>
- Duan H, Lu X, Lian CL, An Y, Xia XL et al (2016) Genome-wide analysis of microRNA responses to the phytohormone abscisic acid in *Populus euphratica*. *Front Plant Sci* 7:1184. <https://doi.org/10.3389/fpls.2016.01184>
- Fan JM (2003) Studies on cyto-morphological and cyto-chemical features of cytoplasmic-nuclear male-sterile lines of soybeans (*Glycine max* (L.) Merr.). M.Sc. thesis, Nanjing Agricultural University, Nanjing, pp 32–41
- Gai JY, Cui ZL, Ji DF, Ren ZJ, Ding DR (1995) A report on the nuclear cytoplasmic male sterility from a cross between two soybean cultivars. *Soy Genet Newsl* 22:55–58
- German MA, Pillay M, Jeong DH, Hetawal A, Luo S et al (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat Biotechnol* 26:941–946. <https://doi.org/10.1038/nbt1417>
- Hölzle A, Jonietz C, Törjek O, Altmann T, Binder S et al (2011) A restorer of fertility-like *PPR* gene is required for 5'-end processing of the nad4 mRNA in mitochondria of *Arabidopsis thaliana*. *Plant J* 65(5):737–744. <https://doi.org/10.1111/j.1365-3113.2010.04460.x>
- Hu J, Wang K, Huang W, Liu G, Gao Y, Wang J et al (2012) The rice pentatricopeptide repeat protein *RF5* restores fertility in Hong-Lian cytoplasmic male-sterile lines via a complex with the glycine rich protein *GRP162*. *Plant Cell* 24:109–122. <https://doi.org/10.1105/tpc.111.093211>
- Iwabuchi M, Koizuka N, Fujimoto H, Sakai T, Imamura J (1999) Identification and expression of the kosena radish (*Raphanus sativus* cv. *Kosena*) homologue of the ogura radish CMS-associated gene, *orf138*. *Plant Mol Biol* 39:183–188. <https://doi.org/10.1023/A:1006198611371>
- Jiang JX, Lv ML, Liang Y, Ma ZM, Cao JS (2014) Identification of novel and conserved miRNAs involved in pollen development in *Brassica campestris* ssp. *chinensis* by high-throughput sequencing and degradome analysis. *BMC Genom* 15:146. <https://doi.org/10.1186/1471-2164-15-146>
- Jordan DR, Mace ES, Henzell RG, Klein PE, Klein RR (2010) Molecular mapping and candidate gene identification of the *Rf2* gene for pollen fertility restoration in sorghum [*Sorghum bicolor* (L.) Moench]. *Theor Appl Genet* 120:1279–1287. <https://doi.org/10.1007/s00122-009-1255-3>
- Kazama T, Toriyama K (2003) A pentatricopeptide repeat-containing gene that promotes the processing of aberrant *atp6* RNA of cytoplasmic male-sterile rice. *FEBS Lett* 544:99–102. [https://doi.org/10.1016/S0014-5793\(03\)00480-0](https://doi.org/10.1016/S0014-5793(03)00480-0)
- Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* 42:D68–D73. <https://doi.org/10.1093/nar/gkt1181>
- Kulcheski FR, Marcelino-Guimaraes FC, Nepomuceno AL, Abdelnoor RV, Margis R (2010) The use of microRNAs as reference genes for quantitative polymerase chain reaction in soybean. *Anal Biochem* 406:185–192. <https://doi.org/10.1016/j.ab.2010.07.020>
- Li JJ, Yang SP, Gai JY (2017a) Transcriptome comparative analysis between the cytoplasmic male sterile line and fertile line in soybean (*Glycine max* (L.) Merr.). *Genes Genomics* 39:1117–1127. <https://doi.org/10.1007/s13258-017-0578-8>
- Li YW, Ding XL, Wang X, He TT, Zhang H et al (2017b) Genome-wide comparative analysis of DNA methylation between soybean cytoplasmic male-sterile line NJCMS5A and its maintainer NJCMS5B. *BMC Genom* 18:596. <https://doi.org/10.1186/s12864-017-3962-5>
- Liu R, Lai B, Hu B, Qin YH, Hu GB et al (2017) Identification of microRNAs and their target genes related to the accumulation of anthocyanins in *Litchi chinensis* by high-throughput sequencing and degradome analysis. *Front Plant Sci* 7:2059. <https://doi.org/10.3389/fpls.2016.02059>
- Liu C, Liu XX, Wu WL, Fu WM, Wang FD et al (2018) Identification of miRNAs and their targets in regulating tuberous root development in radish using small RNA and degradome analyses. *3 Biotech* 8:311. <https://doi.org/10.1007/s13205-018-1330-z>
- Meyers BC, Axtell MJ, Bartel B, Bartel DP, Baulcombe D et al (2008) Criteria for annotation of plant microRNAs. *Plant Cell* 20:3186–3190. <https://doi.org/10.1105/tpc.108.064311>

- Mu JY, Tan HL, Hong SL, Liang Y, Zuo JR (2013) *Arabidopsis* transcription factor genes *NF-YA1*, 5, 6, and 9 play redundant roles in male gametogenesis, embryogenesis, and seed development. *Mol Plant* 6:188–201. <https://doi.org/10.1093/mp/sss061>
- Nasrin Z, Yoshikawa M, Nakamura Y, Begum S, Nakaba S et al (2010) Overexpression of a fungal laccase gene induces nondehiscent anthers and morphological changes in flowers of transgenic tobacco. *J Wood Sci* 56:460–469. <https://doi.org/10.1007/s10086-010-1126-1>
- Nie SS, Xu L, Wang Y, Huang DQ, Muleke EM et al (2015) Identification of bolting-related microRNAs and their targets reveals complex miRNA-mediated flowering-time regulatory networks in radish (*Raphanus sativus* L.). *Sci Rep* 5: 14034. <https://doi.org/10.1038/srep14034>
- Samad AFA, Nazaruiddin N, Murad AMA, Jani J, Zainal Z et al (2018) Deep sequencing and in silico analysis of small RNA library reveals novel miRNA from leaf *Persicaria minor* transcriptome. *3 Biotech* 8:136. <https://doi.org/10.1007/s13205-018-1164-8>
- Schiefthaler U, Balasubramanian S, Sieber P, Chevalier D, Wisman E et al (1999) Molecular analysis of NOZZLE, a gene involved in pattern formation and early sporogenesis during sex organ development in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 96:11664–11669
- Shen YO, Zhang ZM, Lin HJ, Liu HL, Chen J et al (2011) Cytoplasmic male sterility-regulated novel microRNAs from maize. *Funct Integr Genom* 11:179–191. <https://doi.org/10.1007/s10142-010-0202-3>
- Song QX, Liu YF, Hu XY, Zhang WK, Ma B et al (2011) Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing. *BMC Plant Biol* 11:5. <https://doi.org/10.1186/1471-2229-11-5>
- Spanudakis E, Jackson S (2014) The role of microRNAs in the control of flowering time. *J Exp Bot* 65:365–380. <https://doi.org/10.1093/jxb/ert453>
- Tang G (2010) Plant microRNAs: an insight into their gene structures and evolution. *Semin Cell Dev Biol* 21:782–789. <https://doi.org/10.1016/j.semcdb.2010.07.009>
- Wang YM, Ding Y, Liu JY (2016a) Identification and profiling of microRNAs expressed in elongating cotton fibers using small RNA deep sequencing. *Front Plant Sci* 7:1722. <https://doi.org/10.3389/fpls.2016.01722>
- Wang ZS, Wang Y, Kohalmi SE, Amyot L, Hannoufa A (2016b) SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 2 controls floral organ development and plant fertility by activating *ASYMMETRIC LEAVES2* in *Arabidopsis thaliana*. *Plant Mol Biol* 92:661–674. <https://doi.org/10.1007/s11103-016-0536-x>
- Wang DG, Zhang L, Li JK, Hu GY, Wu Q et al (2016c) The restorer gene for soybean M-type cytoplasmic male sterility, *Rf-m*, is located in a *PPR* gene-rich region on chromosome 16. *Plant Breed* 135:342–348. <https://doi.org/10.1111/pbr.12357>
- Wei YY, Chen S, Yang PC, Ma ZY, Kang L (2009) Characterization and comparative profiling of the small RNA transcriptomes in two phases of locust. *Genome Biol* 10:R6. <https://doi.org/10.1186/gb-2009-10-1-r6>
- Wei MM, Wei HL, Wu M, Song MZ, Zhang JF et al (2013) Comparative expression profiling of miRNA during anther development in genetic male sterile and wild type cotton. *BMC Plant Biol* 13:66. <https://doi.org/10.1186/1471-2229-13-66>
- Wei XC, Zhang XH, Yao QJ, Yuan YX, Li XX et al (2015) The miRNAs and their regulatory networks responsible for pollen abortion in Ogura-CMS Chinese cabbage revealed by high-throughput sequencing of miRNAs, degradomes, and transcriptomes. *Front Plant Sci* 6:894. <https://doi.org/10.3389/fpls.2015.00894>
- Xing SP, Salinas M, Höhmann S, Berndtgen R, Huijser P (2010) miR156-targeted and nontargeted SBP-Box transcription factors act in concert to secure male fertility in *Arabidopsis*. *Plant Cell* 22:3935–3950. <https://doi.org/10.1105/tpc.110.079343>
- Yan JJ, Zhang HY, Zheng YZ, Ding Y (2015) Comparative expression profiling of miRNAs between the cytoplasmic male sterile line MeixiangA and its maintainer line MeixiangB during rice anther development. *Planta* 241:109–123. <https://doi.org/10.1007/s00425-014-2167-2>
- Yang WC, Ye D, Xu J, Sundaresan V (1999) The SPOROCTELESS gene of *Arabidopsis* is required for initiation of sporogenesis and encodes a novel nuclear protein. *Genes Dev* 13:2108–2117
- Yang SP, Duan MP, Meng QC, Qiu J, Fan JM et al (2007) Inheritance and gene tagging of male fertility restoration of cytoplasmic-nuclear male-sterile line NJCMS1A in soybean. *Plant Breeding* 126:302–305. <https://doi.org/10.1111/j.1439-0523.2007.01356.x>
- Yang JH, Liu XY, Xu BC, Zhao N, Yang XD et al (2013) Identification of miRNAs and their targets using high-throughput sequencing and degradome analysis in cytoplasmic male-sterile and its maintainer fertile lines of *Brassica juncea*. *BMC Genom* 14:9. <https://doi.org/10.1186/1471-2164-14-9>
- Yu JH, Zhao YX, Qin YT, Yue B, Zheng YL et al (2013) Discovery of microRNAs associated with the S type cytoplasmic male sterility in maize. *J Integr Agric* 12:229–238. [https://doi.org/10.1016/S2095-3119\(13\)60222-1](https://doi.org/10.1016/S2095-3119(13)60222-1)
- Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA (2006) Evidence that miRNAs are different from other RNAs. *Cell Mol Life Sci* 63:246–254. <https://doi.org/10.1007/s00018-005-5467-7>
- Zhang DD, Liu D, Lv XM, Wang Y, Xun ZL et al (2014) The cysteine protease *CEP1*, a key executor involved in tapetal programmed cell death, regulates pollen development in *Arabidopsis*. *Plant Cell* 26:2939–2961. <https://doi.org/10.1105/tpc.114.127282>
- Zhang HY, Hu JH, Qian Q, Chen H, Jin J et al (2016) Small RNA profiles of the rice PTGMS line Wuxiang S reveal miRNAs involved in fertility transition. *Front Plant Sci* 7:514. <https://doi.org/10.3389/fpls.2016.00514>
- Zhao D, Xia X, Wei M, Sun J, Meng J et al (2017) Overexpression of herbaceous peony miR156e-3p improves anthocyanin accumulation in transgenic *Arabidopsis thaliana* lateral branches. *3 Biotech* 7:379. <https://doi.org/10.1007/s13205-017-1011-3>
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406–3415. <https://doi.org/10.1093/nar/gkg595>

Affiliations

Xianlong Ding^{1,2} · Hao Zhang¹ · Hui Ruan¹ · Yanwei Li¹ · Linfeng Chen¹ · Tanliu Wang¹ · Ling Jin¹ · Xiaoqiang Li¹ · Shouping Yang¹ · Junyi Gai¹

Xianlong Ding
xlding2012@163.com

Hao Zhang
haozhang1223@163.com

Hui Ruan
ruanhui613@163.com

Yanwei Li
wwyxiao2013@163.com

Linfeng Chen
linfeng_chen2015@163.com

Tanliu Wang
tanliuWTL@126.com

Ling Jin
msjinling@163.com

¹ Soybean Research Institute, National Center for Soybean Improvement, Key Laboratory of Biology and Genetic Improvement of Soybean (General, Ministry of Agriculture), State Key Laboratory of Crop Genetics and Germplasm Enhancement, Jiangsu Collaborative Innovation Center for Modern Crop Production, College of Agriculture, Nanjing Agricultural University, Nanjing 210095, China

² College of Life Sciences, Nanjing Agricultural University, Nanjing 210095, China