Planta (2021) 253:59
https://doi.org/10.1007/s00425-021-03568-6

ORIGINAL ARTICLE

Identification of microRNAs and their gene targets in cytoplasmic male sterile and fertile maintainer lines of pigeonpea

Abhishek Bohra[1] · Prasad Gandham[2] · Abhishek Rathore[2] · Vivek Thakur[3] · Rachit K. Saxena[2] · S. J. Satheesh Naik[1] · Rajeev K. Varshney[2] · Narendra P. Singh[1]

Received: 25 June 2020 / Accepted: 4 January 2021 / Published online: 4 February 2021

Abstract

Main conclusion **Comparative analysis of genome-wide miRNAs and their gene targets between cytoplasmic male sterile (CMS) and fertile lines of pigeonpea suggests a possible role of miRNA-regulated pathways in reproductive development.**

Abstract Exploitation of hybrid vigor using CMS technology has delivered nearly 50% yield gain in pigeonpea. Among various sterility-inducing cytoplasms (A_1–A_9) reported so far in pigeonpea, A_2 and A_4 are the two major sources that facilitate hybrid seed production. Recent evidence suggests involvement of micro RNA in vast array of biological processes including plant reproductive development. In pigeonpea, information about the miRNAs is insufficient. In view of this, we sequenced six small RNA libraries of CMS line UPAS 120A and isogenic fertile line UPAS 120B using Illumina technology. Results revealed 316 miRNAs including 248 known and 68 novel types. A total of 637 gene targets were predicted for known miRNAs, while 324 genes were associated with novel miRNAs. Degradome analysis revealed 77 gene targets of predicted miRNAs, which included a variety of transcription factors playing key roles in plant reproduction such as F-box family proteins, apetala 2, auxin response factors, ethylene-responsive factors, homeodomain-leucine zipper proteins etc. Differential expression of both known and novel miRNAs implied roles for both conserved as well as species-specific players. We also obtained several miRNA families such as miR156, miR159, miR167 that are known to influence crucial aspects of plant fertility. Gene ontology and pathway level analyses of the target genes showed their possible implications for crucial events during male reproductive development such as tapetal degeneration, pollen wall formation, retrograde signaling etc. To the best of our knowledge, present study is first to combine deep sequencing of small RNA and degradome for elucidating the role of miRNAs in flower and male reproductive development in pigeonpea.

Keywords Male sterility · microRNA · Degradome · Sequencing · Gene expression

Abbreviations

AGO	Argonaute
AP2	Apetala 2
ARF	Auxin response factor
Bp	Base pair
CDS	Coding sequence
CMS	Cytoplasmic male sterility
DCL	Dicer-like
GMUCT	Genome-wide mapping of uncapped transcripts
GO	Gene ontology
GRF	Growth-regulating factor
HD-Zip	Homeodomain-leucine zipper
miRNA	Micro RNA
MYB	Myeloblastosis
nc	Non-coding
NGS	Next generation sequencing
nt	Nucleotide
ORF	Open reading frame
PARE	Parallel analysis of RNA ends

Communicated by Anastasios Melis.

✉ Abhishek Bohra
abhi.omics@gmail.com; Abhishek.Bohra@icar.gov.in

[1] ICAR-Indian Institute of Pulses Research (IIPR), Kanpur, India

[2] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India

[3] Hyderabad Central University (HCU), Hyderabad, India

PNRD	Plant Non-coding RNA Database
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA
Rf	Fertility-restoring
RISC	RNA-induced silencing complex
RLM-RACE	RNA ligase-mediated rapid amplification of cDNA ends
rRNA	Ribosomal RNA
SBP	Squamosa promoter binding protein
Sn RNA	Small nuclear RNA
Sno RNA	Small nucleolar RNA
sRNA	Small RNA
TCP	Teosinte branched 1 cycloidea and PCF
TF	Transcription factor
tRNA	Transfer RNA

Introduction

Cytoplasmic male sterility (CMS) is a maternally inherited inability to produce viable pollen grains and/or male reproductive organs (Chen et al. 2017). Reported across more than 150 plant species, CMS has been proven to be a great tool not only for heterosis utilization but also to elucidate cytoplasmic-nuclear communication (Bohra et al. 2016). Research has shown that CMS results from impaired harmony between the nuclear and mitochondrial genomes (Horn et al. 2014). The CMS phenotypes are rescued following action of fertility-restoring (*Rf*) elements located in the nucleus. Several mitochondrial genes and *Rf* loci responsible for CMS-induction and restoration respectively have been identified in different crops (Bohra et al. 2016; Chen et al. 2017).

Pigeonpea (*Cajanus cajan* L. Millspaugh) is an important grain legume crop widely cultivated in semi-arid tropics due to its higher protein and nutrient content, less input requirement and hardy nature (Bohra et al. 2020). Since the first case of CMS was observed in pigeonpea (Reddy and Faris 1981), a variety of CMS sources have been discovered based on different wild relatives (Bohra et al. 2010; Saxena et al. 2010). Of these, two CMS sources i.e. A_2 and A_4 derived from *Cajanus scarabaeoides* and *Cajanus cajanifolius*, respectively have been widely implicated for efficient hybrid seed production (Bohra et al. 2020). Sequence analysis of mitochondrial genomes in pigeonpea has delineated a set of open reading frames (ORFs) in mitochondria that could influence CMS occurrence (Tuteja et al. 2013). Sinha et al. (2015) have shown association of a 10-bp deletion in the mitochondrial *nad7* gene with male sterility conferred by A_4 cytoplasm. Notwithstanding the discovery of the CMS-related mitochondrial components, the regulatory mechanisms underlying CMS trait still remains poorly understood in pigeonpea. At the same time, preliminary results point to a possibility that male sterility in A_2-CMS may involve mechanisms different than that of A_4-CMS (Sinha et al. 2015; Bohra et al. 2017).

Recent studies have established regulatory roles of non-coding (nc) RNA molecules in a variety of biological processes including plant reproductive development (Mishra and Bohra 2018). Among nc RNAs, microRNAs (miRNA) are the second most important class of endogenous riboregulators (20 to 24-nt) that modulate gene expression through translational repression or RNA degradation. Association of miRNAs with fertility defects/male sterility in plants is evident from overexpression studies of miRNAs such miR167 (Ru et al. 2006) and miR159 (Anthony 2005) in Arabidopsis and rice (Tsuji et al. 2006).

As reviewed by Voinnet (2009), plant miRNA biogenesis involves formation of precursor miRNA (pre-miRNAs) from the action of Dicer-like protein DCL 1 on primary miRNAs (pri-miRNAs) in the nucleus. Pri-miRNAs are transcribed from intergenic regions by RNA polymerase II. Pre-miRNA processing yields an RNA duplex (miRNA/miRNA*), which is transported to cytoplasm by HASTY (the plant homolog of exportin 5). In cytoplasm, one of the duplex strands (mature miRNA/guide strand) is incorporated into RNA-induced silencing complex (RISC) to facilitate repression of the target transcripts.

High-throughput assays based on next generation sequencing (NGS) have allowed deep investigations on small RNA expression. Further, perfect or near perfect complementary base pairing of miRNAs with the target sequence forms the basis of discovery and validation of genes targeted by the identified miRNAs. Experimental validation of the mRNA target is performed through various procedures like AGO protein co-immunoprecipitation and RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) (Chen et al. 2016). To this end, degradome sequencing allows discovery of global gene targets based on large-scale analysis of the mRNA cleavage events directed by miRNAs. Small RNA sequencing in combination with degradome and bioinformatics analyses has greatly enhanced our capacity to elucidate miRNAs and their gene targets.

Deep miRNA profiling of male fertile and sterile plants has been reported recently in different plant species. To date, limited information is available on miRNAs in pigeonpea, which is based on computational analysis of pigeonpea genome (Kompelli et al. 2015; Nithin et al. 2017). A more recent study examined expression change of miRNAs in seedlings of two A_4-based CMS hybrids to demonstrate role of miRNAs towards heterosis in pigeonpea (Sinha et al. 2020). We constructed six small RNA libraries from floral buds of A_2-CMS line (UPAS 120A) and fertile maintainer (UPAS 120B). We identified genome-wide miRNAs from the deep-sequenced small RNA dataset. We further predicted the genes targeted by the identified miRNAs and

carried out their comprehensive functional annotation. We also constructed three degradome libraries to verify the computationally-predicted mRNA targets. To the best of our knowledge, this study is the first to report deep sequencing of small RNA from floral buds of isogenic pigeonpea lines to elucidate genome-wide miRNAs and corresponding biological functions.

Materials and methods

Plant materials, sample collection and total RNA extraction

The CMS line UPAS 120A and maintainer line UPAS 120B were grown at main farm of ICAR-Indian Institute of Pulses Research (IIPR), Kanpur, India (26°27′N latitude, 80°14′E longitude and 152.4 above msl) under normal conditions. The CMS line was developed through standard backcross procedure using A_2-CMS line GT 288A as the donor parent and UPAS 120 (designated as UPAS 120B afterwards), a popular pigeonpea variety, as the recurrent parent (Singh et al. 2009). The pollen fertility assay of the CMS and fertile lines included visual inspection, 1% acetocarmine and 1% iodide potassium iodide (IKI) solution. In the last week of February 2017, the unopened floral buds (9–11 mm size) were collected from the male sterile (UPAS 120A) and fertile (UPAS 120B) lines grown in the field [Temperature: 31.5°/15 °C; relative humidity (%): 82–56; sunshine hours: 10.5; rainfall: 0 mm]. Three biological replicates were collected from each line and used for RNA extraction. Total RNA was isolated from snap-frozen buds with TRIzol reagent (Invitrogen, USA) according to the manufacturer's protocols, and treated with RQ1 RNase-free DNase (Promega, USA). The quality of the total RNA was checked on 1% denaturing gel to check the presence of 28S and 18S ribosomal RNA. Total RNA was quantified using Qubit fluorometer (Invitrogen, USA).

Construction and sequencing of small RNA libraries

Six small RNA libraries were constructed with the Illumina TruSeq small RNA preparation kit from 1 µg of RNA. Initially 3′ and 5′ adaptors were ligated to each end of the RNA molecule and cDNA was created following reverse transcription. cDNA was then PCR amplified using a common primer and an index primer to create cDNA construct. The cDNA construct was then purified using 6% Novex TBE gel. After gel purification, cDNA was extracted and concentrated by ethanol precipitation. Final library was validated by Bio-Analyzer 2100 using high-sensitivity DNA chip (Agilent Technologies, USA). Following the manufacturer's protocol, the small RNA libraries were subjected to deep sequencing using Illumina technology (Xcelris Labs Limited, India).

Computational analysis of sequencing data and identification of miRNAs

Sequencing reads in the fastq format were quality checked using FastQC v. 0.11.8 tool (Andrews 2010). Then Illumina Truseq small RNA were trimmed and low quality reads (> 20% of the bases having phred quality score < 10) were discarded using Trimmomatic v. 0.36 by specifying parameters SLIDINGWINDOW:4:20 and minimum read length cutoff of 16 nt (Bolger et al. 2014). The trimmed reads were again quality checked and aligned by allowing two mismatches to the pigeonpea genome using Bowtie v. 2.3.5.1 (Langmead and Salzberg 2012), and unaligned reads were discarded. Reads corresponding to the coding sequence (CDS), ncRNA sequences downloaded from Plant Non-coding RNA Database (http://structuralbiology.cau.edu.cn/PNRD/) like tRNA, rRNA, small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA) were also excluded from the downstream analysis. Remaining reads were trimmed to the maximum length of 30 nt by using Cutadapt v. 2.4 (Martin 2011).

Identification of known and putative novel miRNAs was performed using mirPRo v. 1.1.4 (Shi et al. 2015) pipeline by enabling read quality filter with 95% of bases for each read with sequencing score higher than 20 (*fastq_quality_filter* -q 20 and -p 95) and miRBase v. 22.1 (Kozomara and Griffiths-Jones 2014) was given as the reference database. Hairpin loop structures of miRNAs were predicted from their pre-miRNA sequences using Mfold v. 3.6 (Zuker 2003). GO analysis was performed to establish their functional roles using WEGO (Ye et al. 2006) and BiNGO (Maere et al. 2005).

Degradome library construction, data analysis and target identification

Three degradome libraries were constructed following manufacturer's protocol and deep sequenced on the Illumina platform (Sandor Life Sciences Private Limited, India). The degradome raw reads were processed using Trimmomatic v. 0.36 to trim adaptors and to remove low-quality reads. The reads aligning to non-coding RNAs were removed and remaining reads were aligned to pigeonpea CDS using Bowtie v. 2.3.5. The reads mapping to the sense strand of transcriptome were given as input to CleaveLand v. 4.5 (Addo-Quaye et al. 2009) to predict miRNA cleavage sites (p value ≤ 0.05). The putative cleavage sites identified were categorized into 0–4 categories based on the read abundance at that position. The categories 0–3 have more than one read mapped at the cleavage site with higher confidence level

prediction, whereas the category 4 has only one read, indicating the minimum confidence. All of the small RNA sequence and degradome sequence data were submitted to the NCBI. The sequencing dataset used in the study is available in the NCBI repository with BioProject: PRJNA625390, BioSamples: SAMN14599215, SAMN14599216, SAMN14599217, SAMN14599218, SAMN14599219, SAMN14599220, SAMN14599221, SAMN14599222, SAMN14599223.

Statistical analysis

To determine differentially expressed miRNAs, count data obtained from mirPRo was provided as an input to DESeq2 Bioconductor package (Love et al. 2014). Three biological replicates each of CMS line (UPAS 120A) and fertile line (UPAS 120B) were used for miRNA identification and target prediction. For each miRNA, DESeq2 fitted negative binomial generalized linear models and Wald test was conducted for significance testing. Count outliers in DESeq2 were detected using Cook's distance and removed from the analysis. The miRNAs showing log2 fold change of ≥ 1 or ≤ 1 and adjusted p value of ≤ 0.01 were considered statistically significant. Respective target genes for known and novel miRNAs were predicted using plant small RNA target analysis server (psRNATarget 2017 release) (Dai et al. 2018) using default parameters. The miRNA target genes were later subjected to GO enrichment analysis with Fisher's exact test from R topGO package (Alexa et al. 2006). The GO terms (from degradome dataset) with p value ≤ 0.01 were considered significantly enriched and represented as word cloud image (Oesper et al. 2011).

Results

Analysis of small RNA libraries from unopened floral buds of UPAS 120A and UPAS 120B

We constructed six independent small RNA libraries from the unopened flower buds of UPAS 120A and UPAS 120B. Sequencing these libraries with Illumina Solexa system yielded a total of raw reads 108701534 and 129714176 for the three samples each of UPAS 120A and UPAS120B, respectively. We obtained a total of 78586006 and 106931476 clean reads for three samples each of UPAS 120A and UPAS 120B, respectively following filtering of low quality reads, adapter/insert contaminants, and tags smaller than 18-nt. Concerning the length distribution of the small RNAs, 24-nt was the most abundant class followed by 21-nt and 22-nt in all six libraries (Fig. 1). Size distribution of all six small RNA libraries is consistent with several reports in other plant species that demonstrated dominance of 24 nt class followed by 21 nt in small RNA transcriptome (Wei et al. 2011; Jiang et al. 2014; Omidvar et al. 2015; Ding et al. 2016; Li et al. 2017). This is a typical range of products (18–24-nt) yielded by the action of Dicer like proteins i. e. DCL1-4 (Voinnet 2009). Sequences corresponding to nc sRNAs including rRNAs, snoRNAs, snRNAs, and

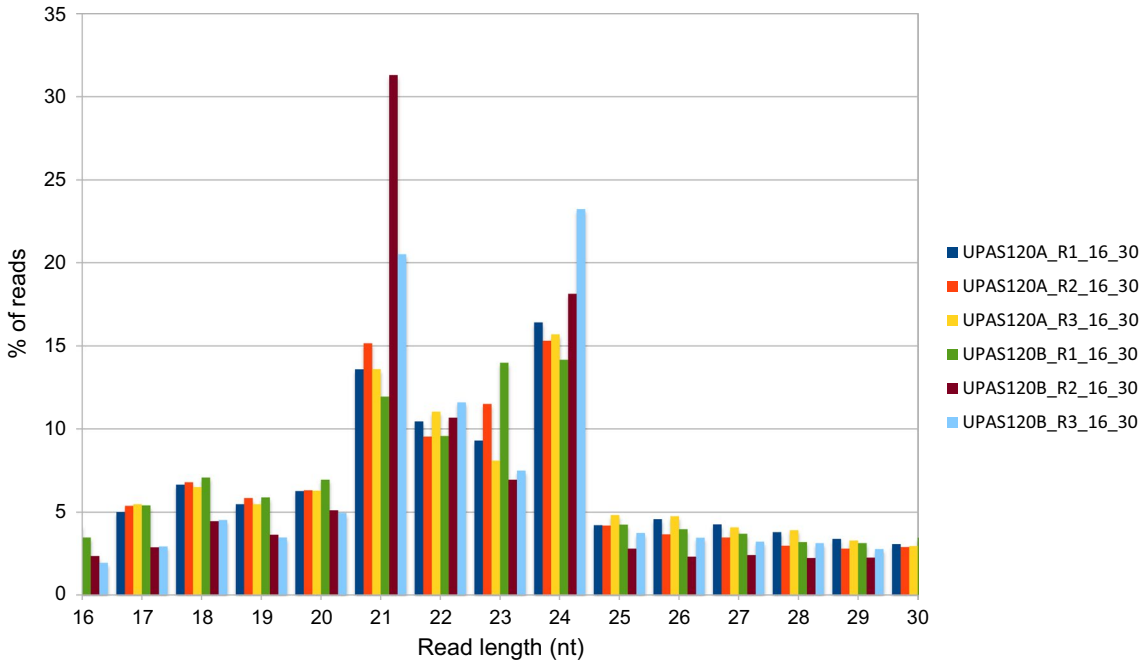


Fig. 1 Size distribution of small RNAs in all six libraries of CMS line (UPAS 120A) and fertile line (UPAS 120B)

tRNAs downloaded from plant non-coding RNA database (http://structuralbiology.cau.edu.cn/PNRD/) were discarded (Table 1).

Identification of known and potential novel miRNAs during anther development

A total of 16,216 known plant miRNAs were used as reference to mirPro, of which 9877 were unique miRNAs. The unique small RNA sequences were aligned to known miRNA precursors/mature miRNA sequences of soybean in mirBase v. 22.1. On the other hand, novel miRNAs involved any sRNAs that found a match in genome sequence, but did not map with the known miRNAs. This approach led to the identification of 316 miRNAs, of which 248 were known and remaining 68 were of novel type (Supplementary Tables 1, 2).

Examples of predicted secondary structures of miRNAs are shown in Fig. 2. Concerning distribution of miRNAs among different families, known miRNAs could be assigned to 46 miRNA families. We noted a highly heterogeneous distribution of miRNAs across different families, with 47.8% of the families having only one or two members (Table 2). The miR166 family (25) had the highest representation followed by miR171 (20) and miR169 (17) families (Fig. 3). Following miRNA categorization of Zhang et al. (2006), we found nine highly conserved families (miR156, miR172, miR171, miR166, miR159, miR396, miR168, miR160 and miR390) and seven moderately conserved families (miR394, miR164, miR169, miR167, miR162, miR398, miR393) in our dataset. Also, four miRNA families obtained here (miR395, miR399, miR403, miR408) were considered as lowly conserved miRNAs. Overall, the length of 21-nt was dominant (64.11%) among the known miRNAs, whereas 24-nt was the most abundant (42.64%) in case of novel miRNAs. Of the 34 families having more than one member, 21 had members with varying length while remaining 13 families contained members with same length. For instance, miRNA166 family had 18 members of 21-nt and seven with 20-nt length. On the other hand, all 20 members of miR171 family had a length of 20-nt (Fig. 3).

We also mapped these miRNAs to the pigeonpea genome using Map2chromosome (Fig. 4a, b). Majority (131) of the miRNAs identified in the present study were located on scaffolds, while 38 and 37 miRNAs were mapped onto CcLG02 and CcLG11, respectively. Two LGs i.e. CcLG05 and CcLG09 did not contain any identified miRNAs (Fig. 4c).

MiRNA target prediction and functional analysis

Unique feature of miRNAs to match perfectly or near-perfectly with the mRNA target sequences is exploited for computational identification of target genes. A variety of

Table 1 Summary of the analysis of high-throughput sequencing reads from six small RNA libraries

Sample name	No. of Raw reads	Trimmomatic		UPAS120* _R*_ trimmo2_cutadapt3.fq.gz		No. of aligned reads to the genome		No. of unaligned Reads to CDS	No. of unaligned Reads to tRNA	No. of unaligned Reads to rRNA	No. of unaligned Reads to snRNA	No. of unaligned Reads to snoRNA	Clean reads between 16 and 30 bp		
		Survived	Dropped	No. of reads	Seq length	Total reads	Unique reads						No. of reads	Unique reads	% reads passed In raw reads
UPAS120A_R1	41,490,235	30,024,457 (99.85%)	43,912 (0.15%)	29,179,541	15–50	28,395,923	1,228,401 (4.32%)	18,680,538	7,170,187	5,699,896	5,139,368	5,100,637	3,452,066	576,492	8.32
UPAS120A_R2	10,457,091	4,102,745 (99.62%)	15,538 (0.38%)	3,313,434	15–50	3,066,059	243,599 (7.94%)	1,808,147	1,013,628	797,625	746,270	743,504	449,363	92,411	4.30
UPAS120A_R3	56,754,208	46,698,032 (99.90%)	47,323 (0.10%)	46,093,031	15–50	43,577,332	2,557,997 (5.87%)	29,469,665	10,728,743	8,281,141	7,280,430	7,232,714	4,593,670	1,112,233	1.96
UPAS120B_R1	13,729,661	7,614,839 (99.79%)	16,152 (0.21%)	7,212,662	15–50	7,024,502	368,459 (5.24%)	4,907,393	1,935,738	1,467,653	1,361,050	1,355,877	774,399	152,942	1.11
UPAS120B_R2	25,222,818	19,370,792 (99.91%)	17,798 (0.09%)	18,475,067	15–50	17,476,849	757,309 (4.33%)	11,189,203	6,084,952	5,198,554	4,862,129	4,848,354	3,315,289	357,330	1.42
UPAS120B_R3	90,761,697	82,138,322 (99.96%)	32,987 (0.04%)	81,243,747	15–50	77,481,106	3,687,284 (4.75%)	55,477,394	22,833,521	19,088,563	17,330,791	17,295,446	10,513,679	1,913,507	2.11

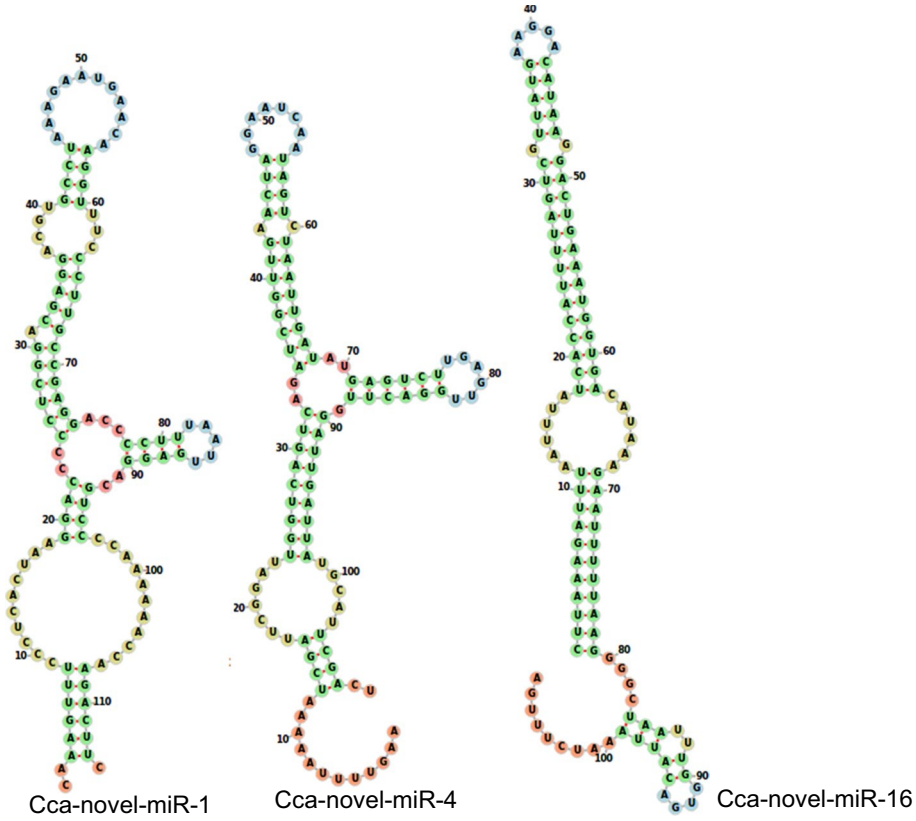


Fig. 2 Predicted secondary structures of novel miRNAs

Table 2 Distribution of miRNAs across various miRNA families in small RNA dataset

Number of members	miRNA families	Number of miRNA families
1	miR1511, miR1514, miR1527, miR2119, miR4371, miR4376, miR4412, miR4414, miR5225, miR5368, miR5786, miR6300	12
2	miR168, miR403, miR828, miR1507, miR1510, miR1513, miR1515, miR2118, miR4416, miR5770	10
3	miR162, miR482, miR4415, miR5037	4
4	miR159, miR398, miR408	3
5	miR530	1
6	miR172, miR2111	2
7	miR160, miR394, miR395	3
8	miR167, miR390	2
9	miR399	1
10	miR393	1
11	miR164	1
14	miR396	1
16	miR156, miR319	2
17	miR169	1
20	miR171	1
25	miR166	1

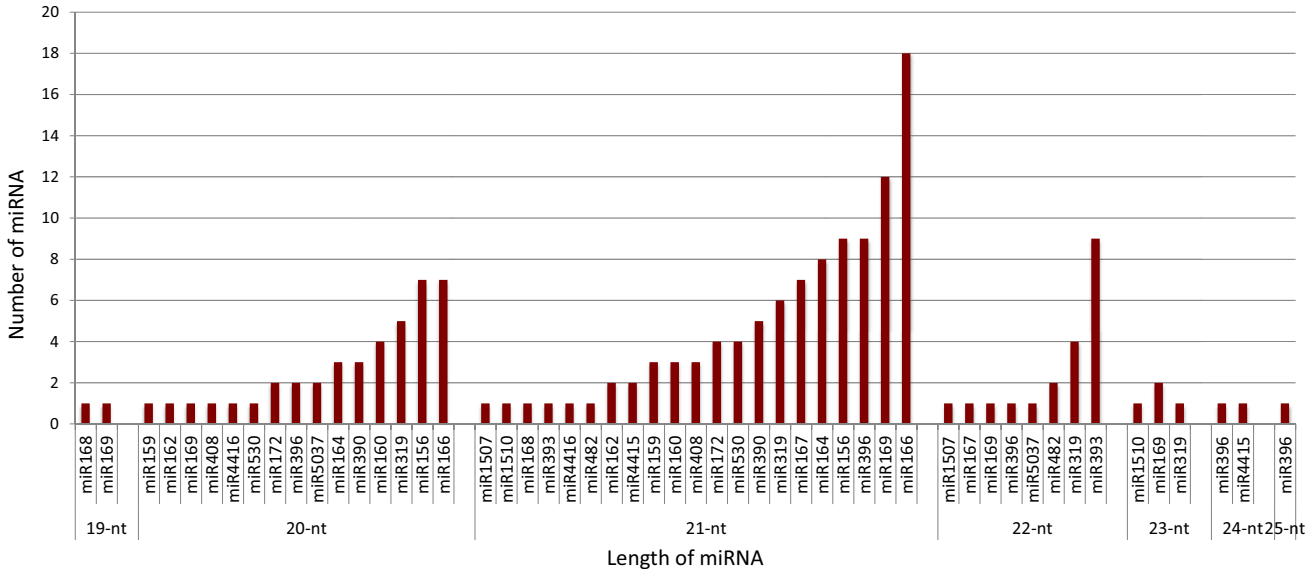


Fig. 3 Known miRNA families identified in the present study and their members of varying lengths

tools such as psRNATarget, TAPIR, miRU, miRNAassist, PatScan, Targetfinder have been development in recent years for miRNA target prediction in plants (Ding et al. 2012). We identified the candidate genes to elucidate biological functions of identified miRNAs. To find the transcripts targeted by these miRNAs, a plant miRNA target prediction web tool psRNAtarget (https://plantgrn.noble.org/psRNATarget/home), was run against the pigeonpea transcripts preloaded in the online tool. A total of 637 gene targets were predicted for known miRNAs, while novel miRNAs targeted 324 genes (Supplementary Tables 3, 4). Figure 4c shows distribution of these genes on pigeonpea chromosomes, with 96 and 70 gene targets located on chromosomes 11 and 3, respectively. A total of 463 gene targets were mapped onto scaffolds.

Results indicate that a single miRNA family can regulate multiple targets. For instance, members of miR171 regulate Scarecrow-like proteins 6, 15, receptor protein kinase-like protein, ribonuclease H, nodulation-signaling pathway 2 protein, Gag/pol polyprotein. On the other hand, most of the members of the same family were found to act on the same target gene.

Target genes of both known and novel miRNAs were assigned to biological processes, molecular functions and cell components based on the GO annotations (Fig. 5). In case of known miRNAs, "Binding" represented the most abundant group in the "molecular function" category followed by "catalytic activity". The most frequent terms in "Biological process" were "Cellular process" and "metabolic process". Likewise, "Cell" and "Cell part" were overrepresented in the "Cellular component" category. GO term classification of genes targeted by novel miRNAs also followed the same pattern.

Similarly, BiNGO analysis revealed "DNA integration", "signalling" and "signal transduction" as the most enriched biological processes (Supplementary Fig. 1). Signalling pathways activated by abscisic acid, auxin etc. were also overrepresented. Importantly, the analysis also demonstrated representation of genes involved in pollen development, pollen tube growth and development, pollen germination, pollination etc. "ATP binding" was found to be the most dominant molecular function, while most of the target genes were found to be associated with "nucleus" followed by "integral component of membrane" and "cytoplasm" among cellular component.

Confirmation of miRNA-guided cleavage of target mRNAs by degradome sequencing

Three independent degradome libraries were constructed from UPAS 120A and UPAS 120B and sequenced. Removal of adaptor sequences and low quality sequences resulted in generation of 18373198, 18161086 and 42490279 clean reads from three degradome libraries. The clean reads thus obtained were mapped to pigeonpea genome. Degradome analysis revealed 77 gene targets ($p = 0.01$), of which 20 were common to psRNATarget (Supplementary Tables 5, 6). Representative target plots of identified miRNA targets are shown in Fig. 6a (miR393k) and 6b (miR172-h-3p). Analysis of the degradome data elucidated a set of gene targets shared between psRNAtarget and degradome analysis, which included MYB, Apetala, F-box family proteins, auxin response

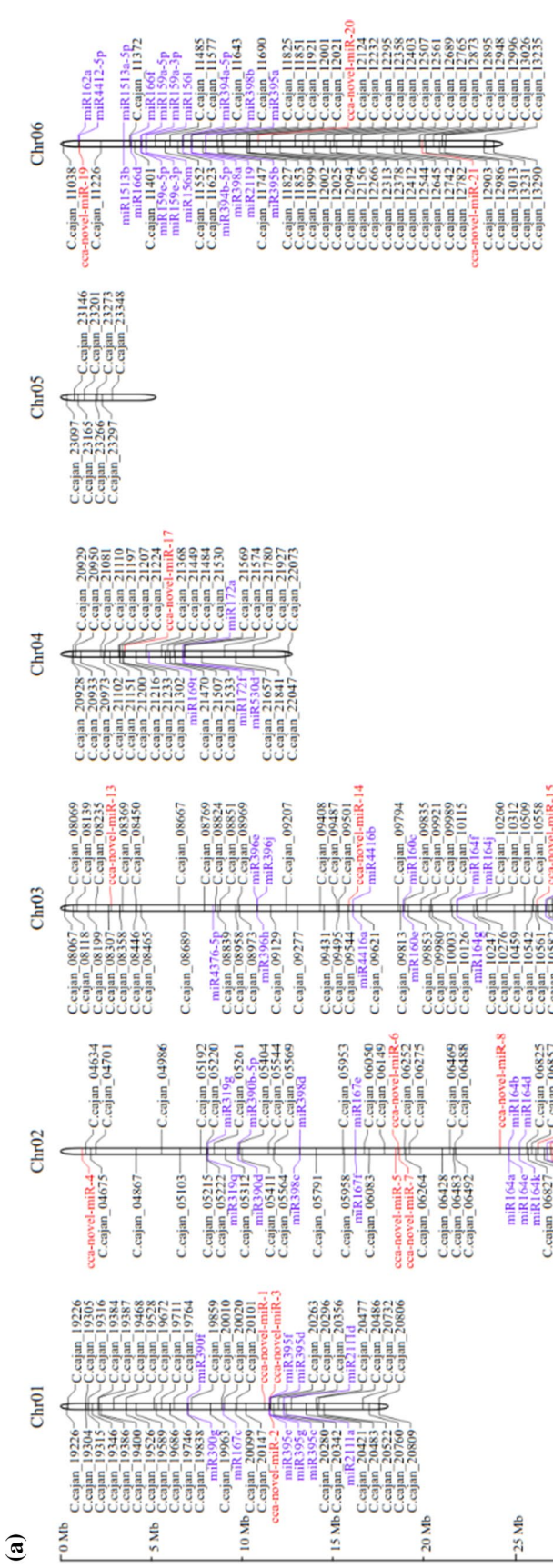
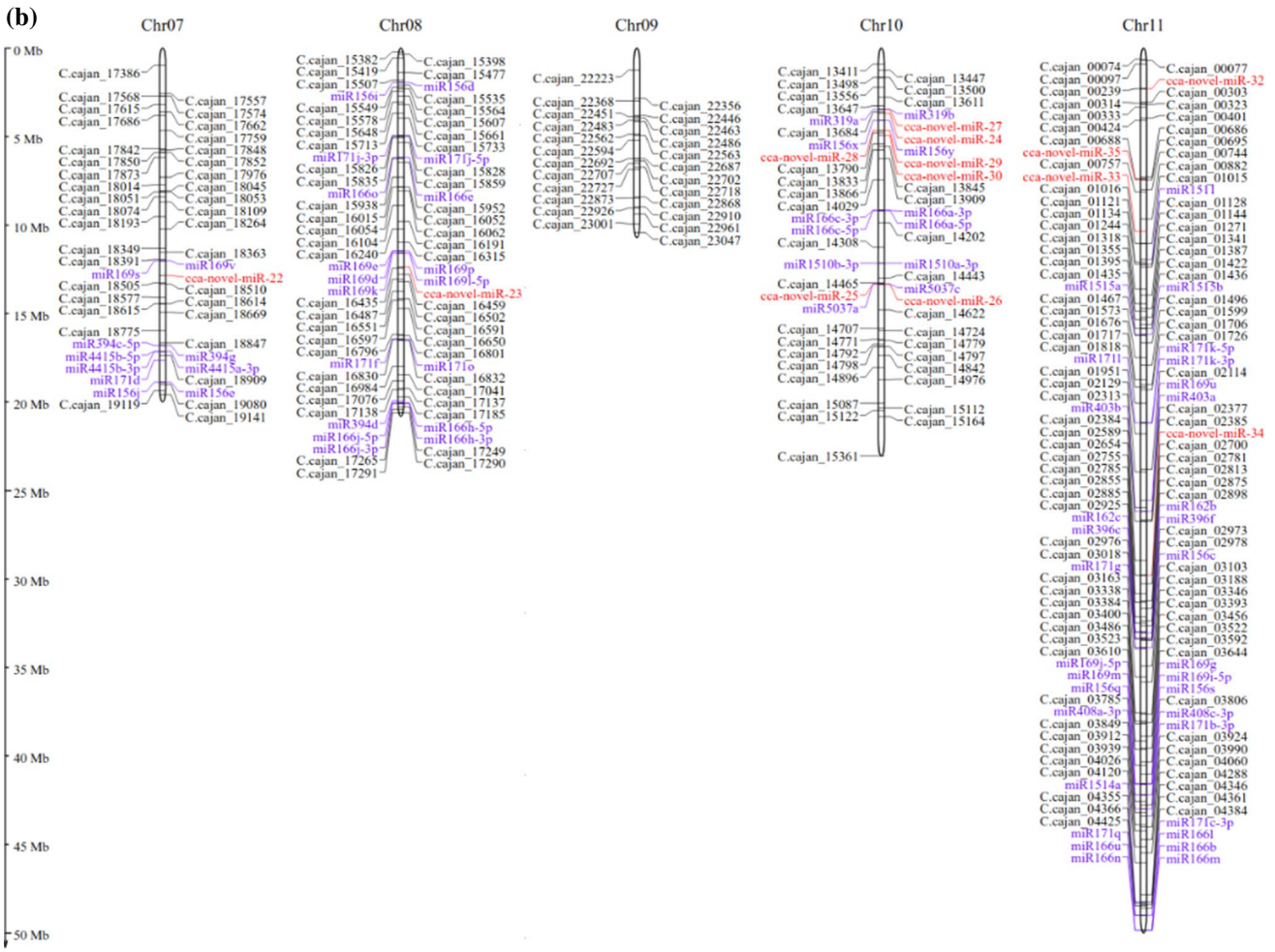


Fig. 4 Chromosomal location of identified miRNAs and target genes. Known miRNAs, novel miRNAs and target genes are shown with blue, red and black colour, respectively. MiRNAs and target genes located on scaffolds are not shown. MiRNAs and target genes on: **a** chromosomes 1–6. **b** Chromosomes 7–11. **c** Distribution patterns of miRNAs and target genes in pigeonpea genome



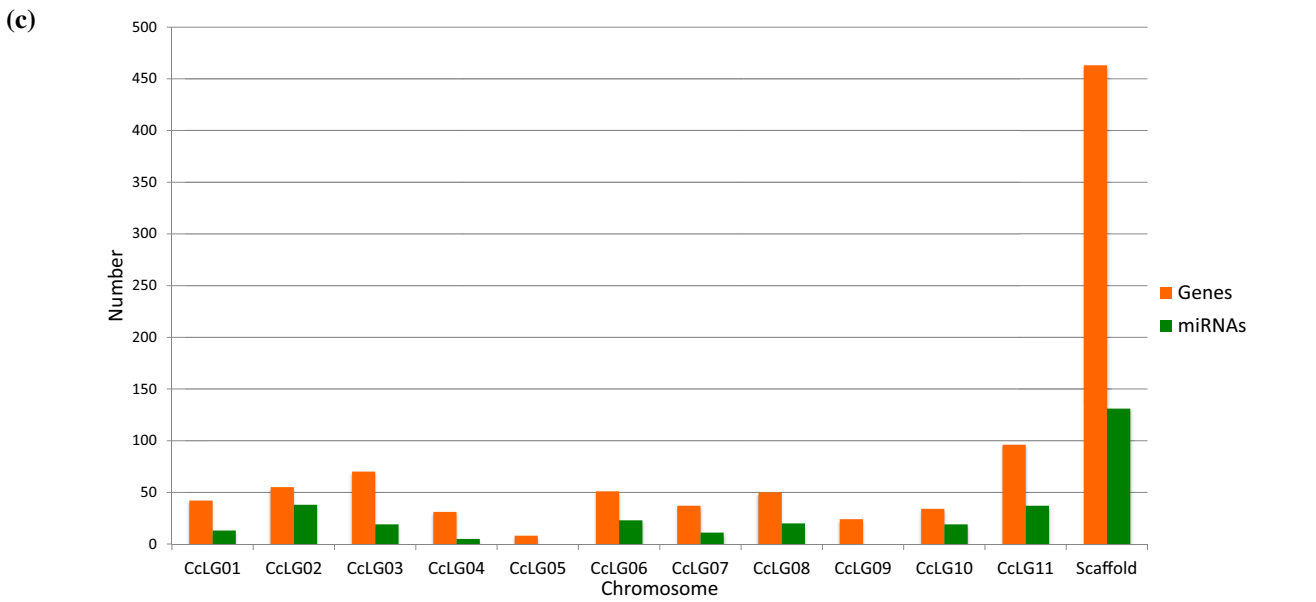


Fig. 4 (continued)

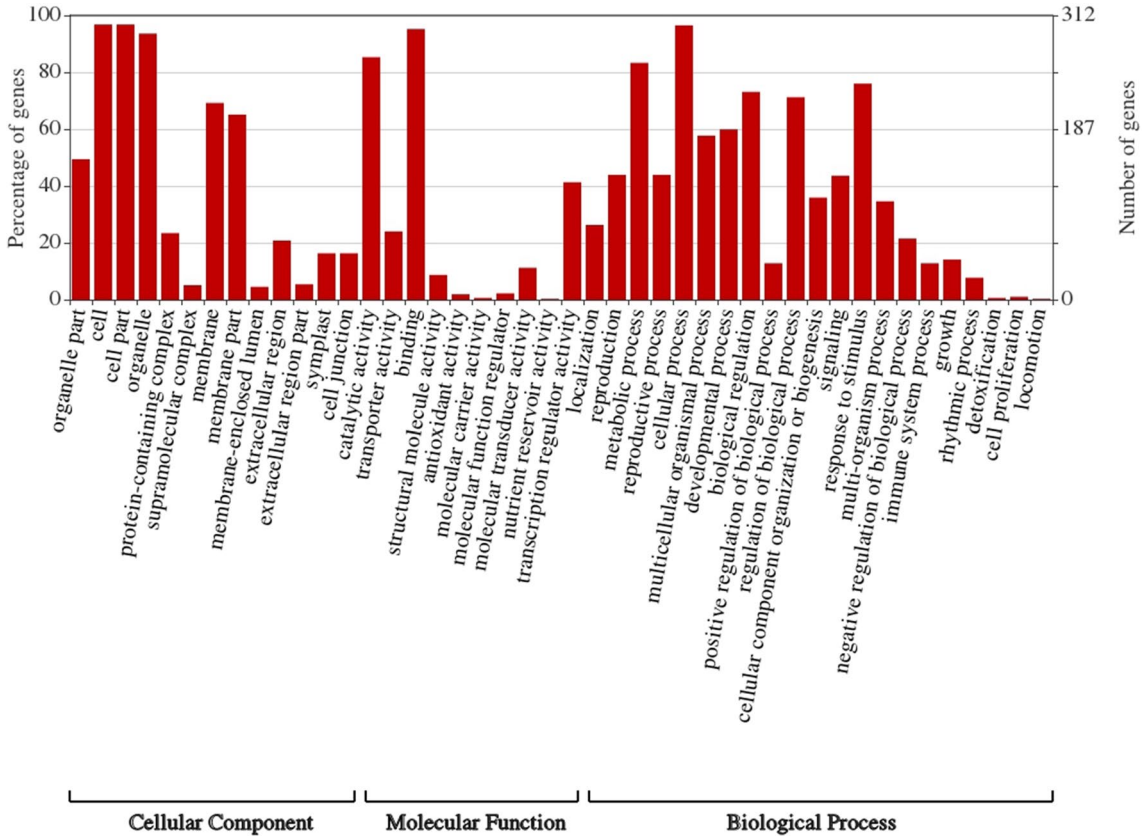


Fig. 5 WEGO plots showing GO terms of target genes predicted for known and novel miRNAs

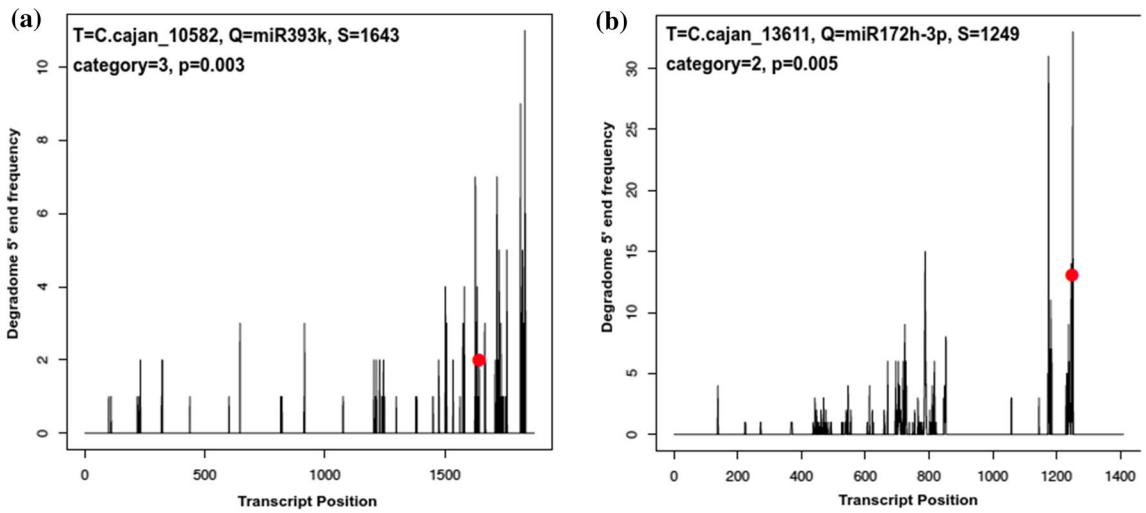


Fig. 6 Representative target plots (*t*-plots) of identified gene targets. Cleavage sites and abundance of raw tags are plotted on *X*-axis and *Y*-axis, respectively. **a** *C.cajan_10582* targeted by miR393k. **b** *C.cajan_13611* targeted by miR172h-3p

factor, TCP 3 transcription factor, ethylene-responsive transcription factor, L-ascorbate oxidase, homeobox-leucine zipper protein, protein transport inhibitor response 1 etc. (Supplementary Table 6). GO analysis of the all degradome target genes led to the identification of 51 GO terms (Supplementary Table 5). Figure 7 shows the most

significant GO terms for genes inferred from the degradome dataset.

Differential expression of miRNAs

Comparison of expression abundance of miRNAs between the CMS line and fertile line provided a set of 28 differentially expressing miRNAs (Table 3, Fig. 8). Interestingly, all members of mir166 family showed upregulation in the fertile line, whereas members of mir171 family were downregulated in the fertile line. Higher expression was also noted for miR 403 in fertile line, while expression levels of six of the novel miRNAs (Cca-novel-miR-4, Cca-novel-miR-31, Cca-novel-miR-32, Cca-novel-miR-43, Cca-novel-miR-49, Cca-novel-miR-52) were supressed in the fertile line.

Discussion

MiRNAs are reported to play crucial roles in a variety of biological processes including floral organ development, hormonal pathways and response to biotic and abiotic stresses (Li et al. 2017). MiRNAs negatively control target genes at post-transcription level through mRNA cleavage and translational repression (Mishra and Bohra 2018), however, researchers have also found cases where miRNAs act as activators (Ding et al. 2012). Methods for miRNA identification and establishment of their function roles have evolved constantly in both animals and plants since the miRNAs were first discovered in nematode *Caenorhabditis elegans* (Wightman et al. 1993) and Arabidopsis (Reinhart et al. 2002). In recent years, high-throughput sequencing methods based on NGS have enabled genome-wide characterization of miRNAs in various plant species.

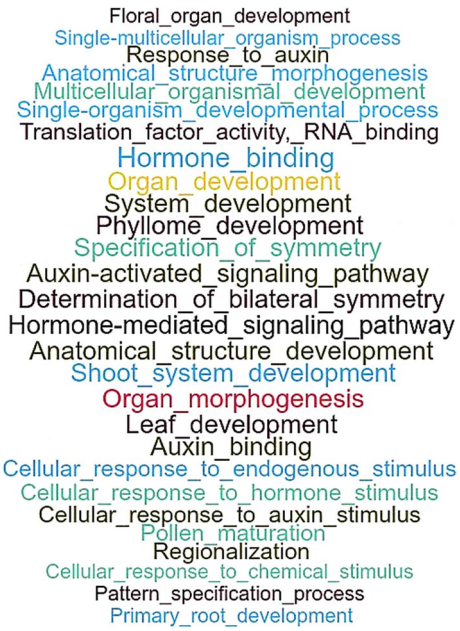


Fig. 7 Word cloud of the enriched GO terms of the target genes from degradome dataset. The font sizes in the word cloud are proportional to the significance of the corresponding GO terms

A growing body of literature indicates that miRNAs orchestrate gene expression and pathways that are involved in CMS induction in plants (Mishra and Bohra 2018). However, only limited reports are currently available in pigeonpea on discovery of miRNAs and their functional roles (Kompelli et al. 2015; Nithin et al. 2017; Sinha et al. 2020). The numbers of miRNA families (46) in the present dataset and members within each family (1–25) were consistent with a previous study in pigeonpea (Kompelli et al. 2015). Higher representation of families like miR166, miR171 and miR159 is evident in pigeonpea genome (Nithin et al. 2017). MiR 169, miR166, miR 156 had the higher number of members in small RNA dataset in rice, Arabidopsis (Zhang et al. 2006) and soybean (Ding et al. 2016).

The gene targets predicted in our study concur with the published reports. For instance, miR166, miR171, miR156, miR159, miR160/miR167 are reported to regulate a variety of transcription factors like HD-Zip, scarecrow-like, squamosa promoter binding protein (SBP)-like (SPL), MYB, ARF, respectively. Consistent with the present study, members of the same miRNA family were found to target different genes like miR166 members targeted AtHB-14 and MLO3 (Sinha et al. 2020).

MiRNA-mediated regulation of several TFs supports their key roles in gene regulatory networks underlying reproductive development in plants. For instance, scarecrow-like and MYB TFs are involved in gibberellic acid signaling, (SBP)-like genes are pivotal to anther development (Lin et al. 2020), whereas ARFs contribute toward anther dehiscence and pollen maturation (Cecchetti et al. 2008) in plants. A recent study in turnip based on small RNA sequencing of floral buds from CMS and maintainer line also revealed HD-Zip, SBP-like genes, MYB, ARFs, AP2, GRF as the potential targets of identified miRNAs (Lin et al. 2020).

Validation of the identified miRNA targets was conventionally performed using target-specific methods such as 5′-RACE, which are less efficient in terms of large-scale target identification (Thomson et al. 2011). Massively parallel sequencing methods like parallel analysis of RNA ends (PARE) or degradome sequencing or genome-wide mapping of uncapped transcripts (GMUCT) are now increasingly employed to allow large-scale identification of miRNA-directed mRNA cleavage events (Addo-Quaye et al. 2008; German et al. 2008). Degradome sequencing has been successfully applied for this purpose across a wide range of plant species (Ding et al. 2012). Degradome analysis confirmed miRNA-mediated regulation of various genes/TFs

Table 3 List of differentially expressed miRNAs between CMS line UPAS 120A and UPAS 120B

S. No	miRNA	Mature miRNA sequence	log2FoldChange	*p*-value
1	cca-novel-miR-4	GAUUGGUCAGUCAGAUCGGU	− 4.358792496	0.000215147
2	cca-novel-miR-31	GCGGAUAUCUUUGAUAAGGUUGAUA	− 2.56042133	0.003383255
3	cca-novel-miR-32	GCGGAUAUCUUUGAUAAGGUUGAUA	− 2.56042133	0.003383255
4	cca-novel-miR-52	ACAUAGUUCUUUUAACACACU	− 3.033640052	0.005688009
5	miR171j-3p	UGAUUGAGCCGUGCCAAUAUC	− 1.672628904	0.008489387
6	cca-novel-miR-43	CCAUGGUUUGGGCAAAAGCU	− 3.491700278	0.009272924
7	miR171e	UGAUUGAGCCGUGCCAAUAUC	− 1.734486281	0.013546884
8	miR403a	UUAGAUUCACGCACAAACUUG	1.896086681	0.01368899
9	miR403b	UUAGAUUCACGCACAAACUUG	1.896086681	0.01368899
10	miR171f	UGAUUGAGCCGUGCCAAUAUC	− 1.66489348	0.020112579
11	miR171g	UGAUUGAGCCGUGCCAAUAUC	− 1.66489348	0.020112579
12	miR166m	CGGACCAGGCUUCAUUCCCC	1.483757874	0.034401074
13	miR171u	UGAUUGAGCCGUGCCAAUAUC	− 1.455077979	0.041309794
14	miR166b	UCGGACCAGGCUUCAUUCCCC	1.45291916	0.042998928
15	miR166n	UCGGACCAGGCUUCAUUCCCC	1.452912553	0.043000073
16	miR166o	UCGGACCAGGCUUCAUUCCCC	1.452880911	0.043002943
17	miR166a-3p	UCGGACCAGGCUUCAUUCCCC	1.452405785	0.043050458
18	miR166c-3p	UCGGACCAGGCUUCAUUCCCC	1.452405785	0.043050458
19	miR166e	UCGGACCAGGCUUCAUUCCCC	1.45238017	0.04306657
20	miR166g	UCGGACCAGGCUUCAUUCCCC	1.452373564	0.043067717
21	miR166i-3p	UCGGACCAGGCUUCAUUCCCC	1.451850845	0.043119225
22	miR166q	UCGGACCAGGCUUCAUUCCC	1.450416344	0.043146756
23	miR166d	UCGGACCAGGCUUCAUUCCCC	1.451867371	0.043166479
24	miR166f	UCGGACCAGGCUUCAUUCCCC	1.451867371	0.043166479
25	miR166s	UCGGACCAGGCUUCAUUCCC	1.449758027	0.043227088
26	miR166t	UCGGACCAGGCUUCAUUCCC	1.449758027	0.043227088
27	miR166p	UCGGACCAGGCUUCAUUCCC	1.449345509	0.04328231
28	cca-novel-miR-49	AGAGGAUUCUUGCUCACGUCGGGC	1.622453943	0.04991182

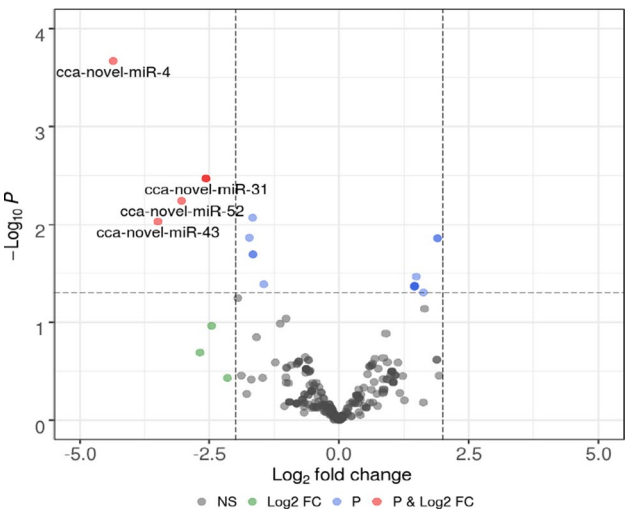


Fig. 8 Volcano plot showing differential miRNA expression between UPAS 120A and UPAS 120B. Upregulated miRNAs are shown on the right side of the plot, whereas left side contains downregulated miRNAs. Statistical significance of miRNA expression level follows the following order: red > blue > green

like F-box family proteins, ARFs, TCP 3 transcription factor, ethylene-responsive transcription factor, L-ascorbate oxidase, HD-Zip, TRANSPORT INHIBITOR RESPONSE 1 etc. that are known to participate in floral and reproductive development (Fang et al. 2016). Yang et al. (2013) made similar observations in *Brassica juncea* following analysis of small RNA and degradome datasets of CMS and fertile lines. As demonstrated in Arabidopsis, TCP transcription factors regulated cytoplasmic-nuclear communication responsible for another-specific expression of several genes (Welchen and Gonzalez 2005).

Differential expression of miRNAs has been associated with occurrence of male sterility in different plant species including rice (Yan et al. 2015), maize (Shen et al. 2011; Yu et al. 2013), soybean (Ding et al. 2016), pepper (Asha et al. 2016; Zhang et al. 2020), cotton (Wei et al. 2013; Zhang et al. 2018; Yu et al. 2020), tomato (Omidvar et al. 2015), radish (Zhang et al. 2016a), *Brassica* (Yang et al. 2013; Jiang et al. 2014; Lin et al. 2020), pummelo (Fang et al. 2016). In our study, a comparative analysis of miRNA expression between CMS and fertile line revealed a significant change

in expression levels of 28 mRNAs. Of these, roles of miR166 and miR171 families in anther and pollen development are well established. MiR166 is among highly conserved miRNA families in plants, and Wei et al. (2011) established MiR166 as one of the "integral regulators" of pollen and sporophyte development based on deep sequencing of small RNAs from sporophytic tissues and pollen in rice. Abundant expression of miR166 during pollen development in both diploid and autotetraploid rice suggested its "conserved" and "essential" contributions toward pollen development (Li et al. 2016). Gene expression modulation by miR166 could be related to anther development and male sterility in tomato (Omidvar et al. 2015). Similarly, influence of miRNA171 silencing in anthers included delayed tapetum degeneration and reduced callose deposition leading to impaired pollen development in tomato (Kravchik et al. 2019). Differential expression of miR171 and corresponding target genes (Scarecrow-like gene and ankyrin repeat domain-containing protein) is reported to modulate fertility transitions of photothermosensitive genic male sterile lines in wheat (Bai et al. 2017) and rice (Zhang et al. 2016b). Another DE miRNA, miR403 is reported to participate in miRNA biogenesis via targeting AGO 2 protein (Fang et al. 2016). Similar to a previous study by Ujino-Ihara et al. (2018), we also obtained several miRNAs families that are key to plant fertility such as miR156, miR159, miR167, miR172, miR319, miR396, miR2118; however, their expression levels between CMS and fertile line did not differ significantly.

Conclusion

High-throughput sequencing of small RNA and degradome libraries from floral buds of isogenic lines was performed for the first time in pigeonpea. A total of 248 known and 68 novel miRNAs were obtained. We predicted that these 316 miRNAs regulate expression of 961 genes engaged in various aspects of flowering and reproductive development in plants. Differential expression between the CMS line and fertile line was obtained for 28 miRNAs. Members of miR166, miR171 and miR403 families and six novel miRNAs (Cca-novel-miR-4, Cca-novel-miR-31, Cca-novel-miR-32, Cca-novel-miR-43, Cca-novel-miR-49, Cca-novel-miR-52) responded differentially to CMS. In summary, the current study contributes new insights on the potential role of miRNA biogenesis in flower development and CMS occurrence in pigeonpea.

Author contribution statement AB conceived the idea, and performed experiments. AR, PG, VT, RKS performed data analysis. RKV, RKS, SJSN and NPS contributed to interpretation of results. AB wrote and revised the manuscript together with PG, AR, RKS, RKV and SNSJ. All authors read and approved the final manuscript.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00425-021-03568-6.

Acknowledgements AB acknowledges financial support from Indian Council of Agricultural Research (ICAR), New Delhi (Grant no.: AGENIASRICOP201501000047).

Compliance with ethical standards

Conflict of interests The author(s) declare that they have no competing interests.

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

Addo-Quaye C, Miller W, Axtell MJ (2009) CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics 25:130–131

Alexa A, Rahnenführer J, Lengauer T (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22:1600–1607

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Anthony AMFG (2005) The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. Plant Cell 17:705–721

Asha S, Sreekumar S, Soniya EV (2016) Unravelling the complexity of microRNA-mediated gene regulation in black pepper (*Piper nigrum* L.) using high-throughput small RNA profiling. Plant Cell Rep 35:53–63

Bai JF, Wang YK, Wang P et al (2017) Uncovering male fertility transition responsive mirna in a wheat photo-thermosensitive genic male sterile line by deep sequencing and degradome analysis. Front Plant Sci 8:1370

Bohra A, Mallikarjuna N, Saxena KB, Upadhyaya H, Vales I, Varshney R (2010) Harnessing the potential of crop wild relatives through genomics tools for pigeonpea improvement. J Plant Biol 37:83–98

Bohra A, Jha UC, Premkumar A, Bisht D, Singh NP (2016) Cytoplasmic male sterility (CMS) in hybrid breeding in field crops. Plant Cell Rep 35:967–993

Bohra A, Jha A, Singh IP, Pandey G, Pareek S, Basu PS, Chaturvedi SK, Singh NP (2017) Novel CMS lines in pigeonpea [*Cajanus cajan* (L.) Millspaugh] derived from cytoplasmic substitutions, their effective restoration and deployment in hybrid breeding. Crop J 5:89–94

Bohra A, Saxena KB, Varshney RK, Saxena RK (2020) Genomics assisted breeding for pigeonpea improvement. Theor Appl Genet 133:1721–1737

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics 15:2114–2120

Cecchetti V, Altamura MM, Falasca G, Costantino P, Cardarelli M (2008) Auxin regulates Arabidopsis anther dehiscence, pollen maturation, and filament elongation. Plant Cell 20:1760–1774

Chen J, Zheng Y, Qin L, Wang Y, Chen L, He Y, Fei Z, Lu G (2016) Identification of miRNAs and their targets through high-throughput sequencing and degradome analysis in male and female *Asparagus officinalis*. BMC Plant Biol 16:80

Chen Z, Zhao N, Li S et al (2017) Plant mitochondrial genome evolution and cytoplasmic male sterility. Crit Rev Plant Sci 36:55–69

Dai X, Zhuang Z, Zhao PX (2018) psRNATarget: a plant small RNA target analysis server (2017 release). Nucleic Acids Res 46(W1):W49–W54

Ding J, Zhou S, Guan J (2012) Finding MicroRNA targets in plants: current status and perspectives. Genom Proteom Bioinform 10:264–275

Ding X, Li J, Zhang H 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

Fang YN, Zheng BB, Wang L, Yang W, Wu XM, Xu Q, Guo WW (2016) High-throughput sequencing and degradome analysis reveal altered expression of miRNAs and their targets in a male-sterile cybrid pummelo (*Citrus grandis*). BMC Genomics 17:591

German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P et al (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 26:941–946

Horn R, Gupta KJ, Colombo N (2014) Mitochondrion role in molecular basis of cytoplasmic male sterility. Mitochondrion 19:198–205

Jiang J, Lv M, Liang Y, Ma Z, Cao J (2014) Identification of novel and conserved miRNAs involved in pollen development in *Brassica campestris* ssp. *chinensis* by high-throughput sequencing and degradome analysis. BMC Genomics 15:146

Kompelli SK, Kompelli VSP, Enjala C, Suravajhala P (2015) Genome-wide identification of miRNAs in pigeonpea (*Cajanus cajan* L.). Aust J Crop Sci 9:215–222

Kozomara A, Griffiths-Jones S (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res 42(1):D68-73

Kravchik M, Stav R, Belausov E, Arazi T (2019) Functional characterization of microRNA171 family in tomato. Plants 8:10

Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359

Li X, Shahid MQ, Wu J, Wang L, Liu X, Lu Y (2016) Comparative small RNA analysis of pollen development in autotetraploid and diploid Rice. Int J Mol Sci 17:499

Li W, He Z, Zhang L, Lu Z, Xu J, Cui J, Wang L, Jin B (2017) miRNAs involved in the development and differentiation of fertile and sterile flowers in *Viburnum macrocephalum f. keteleeri*. BMC Genomics 18:783

Lin S, Su S, Jin L et al (2020) Identification of microRNAs and their targets in inflorescences of an Ogura-type cytoplasmic male-sterile line and its maintainer fertile line of turnip (*Brassica rapa* ssp. *rapifera*) via high-throughput sequencing and degradome analysis. PLoS ONE 15:0236829

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550

Maere S, Heymans K, Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. Bioinformatics 21:3448–3449

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12

Mishra A, Bohra A (2018) Non-coding RNAs and plant male sterility: current knowledge and future prospects. Plant Cell Rep 37:177–191

Nithin C, Thomas A, Basak J, Bahadur RP (2017) Genome-wide identification of miRNAs and lncRNAs in *Cajanus cajan*. BMC Genomics 18:1–14

Oesper L, Merico D, Isserlin R, Bader GD (2011) WordCloud: a Cytoscape plugin to create a visual semantic summary of networks. Source Code Biol Med 6:7

Omidvar V, Mohorianu I, Dalmay T, Fellner M (2015) Identification of miRNAs with potential roles in regulation of anther development and male-sterility in 7B–1 male-sterile tomato mutant. BMC Genomics 16:878

Reddy LJ, Faris DG (1981) A cytoplasmic-genetic male-sterile line in pigeonpea. Int Pigeonpea Newsl 1:16–17

Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP (2002) MicroRNAs in plants. Genes Dev 16:1616–1626

Ru P, Xu L, Ma H, Huang H (2006) Plant fertility defects induced by the enhanced expression of microRNA167. Cell Res 16:457–465

Saxena KB, Sultana R, Mallikarjuna N, Saxena RK, Kumar RV, Sawargaonkar KL (2010) Male-sterility systems in pigeonpea and their role in enhancing yield. Plant Breed 129:125–134

Shen Y, Zhang Z, Lin H et al (2011) Cytoplasmic male sterility regulated novel microRNAs from maize. Funct Integr Genomics 11:179–191

Shi J, Dong M, Li L et al (2015) mirPRo - a novel standalone program for differential expression and variation analysis of miRNAs. Sci Rep 5:14617

Singh IP, Singh BB, Ali I, Kumar S (2009) Diversification and evaluation of cytoplasmic nuclear male sterility system in pigeonpea (*Cajanus cajan*). Indian J Agric Sci 79:291–294

Sinha P, Saxena KB, Saxena RK et al (2015) Association of nad7a gene with cytoplasmic male sterility in pigeonpea. Plant Genome 8:2

Sinha P, Singh VK, Saxena RK, Kale SM, Li Y, Garg V, Meifang T, Khan AW, Kim KD, Chitikineni A, Saxena KB, Sameer Kumar CV, Liu X, Xu X, Jackson S, Powell W, Nevo E, Searle IR, Lodha M, Varshney RK (2020) Genome-wide analysis of epigenetic and transcriptional changes associated with heterosis in pigeonpea. Plant Biotechnol J. https://doi.org/10.1111/pbi.13333

Thomson DW, Bracken CP, Goodall GJ (2011) Experimental strategies for microRNA target identification. Nucleic Acids Res 39:6845–6853

Tsuji H, Aya K, Ueguchi-Tanaka M et al (2006) GAMYB controls different sets of genes and is differentially regulated by microRNA in aleurone cells and anthers. Plant J 47:427–444

Tuteja R, Saxena RK, Davila J, Shah T, Chen W, Xiao YL, Fan G, Saxena KB, Alverson AJ, Spillane C, Town C, Varshney RK (2013) Cytoplasmic male sterility-associated chimeric open reading frames identified by mitochondrial genome sequencing of four *Cajanus* genotypes. DNA Res 20:485–495

Ujino-Ihara T, Ueno S, Uchiyama K, Futamura N (2018) Comprehensive analysis of small RNAs expressed in developing male strobili of *Cryptomeria japonica*. PLoS ONE 13:e0193665

Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136:669–687

Wei LQ, Yan LF, Wang T (2011) Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of *Oryza sativa*. Genome Biol 12:R53

Wei M, Wei H, Wu M, Song M, Zhang J, Yu J, Fan S, Yu S (2013) Comparative expression profiling of miRNA during anther development in genetic male sterile and wild type cotton. BMC Plant Biol 13:66

Welchen E, Gonzalez DH (2005) Differential expression of the Arabidopsis cytochrome c genes Cytc-1 and Cytc-2 evidence for the involvement of TCP-domain protein binding elements in anther- and meristem-specific expression of the Cytc-1 gene. Plant Physiol 139:88–100

Wightman B, Ha I, Ruvkun G (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. Cell 75:855–862

Yan J, Zhang H, Zheng Y, 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

Yang J, Liu X, Xu B, Zhao N, Yang X, Zhang M (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

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J (2006) WEGO: a web tool for plotting GO annotations. Nucleic Acids Res 34(1):W293–W297

Yu JH, Zhao YX, Qin YT, Xiao HL (2013) Discovery of MicroRNAs associated with the S type cytoplasmic male sterility in maize. J Integr Agr 12:229–238

Yu D, Li L, Wei H, Yu S (2020) Identification and profiling of microRNAs and differentially expressed genes during anther development between a genetic male-sterile mutant and its wild type cotton via high-throughput RNA sequencing. Mol Genet Genomics 295:645–660

Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA (2006) Conservation and divergence of plant microRNA genes. Plant J 46:243–259

Zhang W, Xie Y, Xu L et al (2016a) Identification of microRNAs and their target genes explores miRNA mediated regulatory network of cytoplasmic male sterility occurrence during anther development in radish (*Raphanus sativus* L.). Front Plant Sci 7:1054

Zhang H, Hu J, Qian Q, Chen H, Jin J, Ding Y (2016b) Small RNA profiles of the rice PTGMS line Wuxiang S reveal miRNAs involved in fertility transition. Front Plant Sci 7:514

Zhang B, Zhang X, Liu G et al (2018) A combined small RNA and transcriptome sequencing analysis reveal regulatory roles of miRNAs during anther development of upland cotton carrying cytoplasmic male sterile *Gossypium harknessii* (D2) cytoplasm. BMC Plant Biol 18:242

Zhang H, Huang S, Tan J, Chen X, Zhang M (2020) MiRNAs profiling and degradome sequencing between the CMS-line N816S and its maintainer line Ning5m during anther development in pepper (*Capsicum annuum* L.). bioRxiv 1:1

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acid Res 31:3406–3415