Research paper

Novel and conserved microRNAs in soybean floral whorls

F.R. Kulcheski [a,b], L.G. Molina [a,c], G.C. da Fonseca [a], G.L. de Morais [a,d], L.F.V. de Oliveira [a,c], R. Margis [a,b,c,*]

[a] PPGBCM, Centro de Biotecnologia, Laboratório de Genomas e Populações de Plantas, Universidade Federal do Rio Grande do Sul — UFRGS, Porto Alegre, RS, Brazil
[b] Departamento de Biofísica, Universidade Federal do Rio Grande do Sul — UFRGS, Porto Alegre, RS, Brazil
[c] PPGGBM, Universidade Federal do Rio Grande do Sul — UFRGS, Porto Alegre, RS, Brazil
[d] LNCC, Laboratorio Nacional de Ciência da Computação, Petrópolis, RJ, Brazil

ABSTRACT

MicroRNAs (miRNAs) correspond to a class of endogenous small non-coding RNAs (19–24 nt) that regulates the gene expression, through mRNA target cleavage or translation inhibition. In plants, miRNAs have been shown to play pivotal roles in a wide variety of metabolic and biological processes like plant growth, development, and response to biotic and abiotic stress. Soybean is one of the most important crops worldwide, due to the production of oil and its high protein content. The reproductive phase is considered the most important for soybean yield, which is mainly intended to produce the grains. The identification of miRNAs is not yet saturated in soybean, and there are no studies linking them to the different floral organs. In this study, three different mature soybean floral whorls were used in the construction of sRNA libraries. The sequencing of petal, carpel and stamen libraries generated a total of 10,165,661 sequences. Subsequent analyses identified 200 miRNAs sequences, among which, 41 were novel miRNAs, 80 were conserved soybean miRNAs, 31 were new antisense conserved soybean miRNAs and 46 were soybean miRNAs isoforms. We also found a new miRNA conserved in other plant species, and finally one miRNA-sibling of a soybean conserved miRNA. Conserved and novel miRNAs were evaluated by RT-qPCR. We observed a differential expression across the three whorls for six miRNAs. Computational predicted targets for miRNAs analyzed by RT-qPCR were identified and present functions related to reproductive process in plants. In summary, the increased accumulation of specific and novel miRNAs in different whorls indicates that miRNAs are an important part of the regulatory network in soybean flower.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

MicroRNAs (miRNAs) are a class of endogenous small non-coding RNAs (19–24 nt) that regulates gene expression, mainly through mRNA targets cleavage and translation inhibition (Bartel, 2009). Since their recent discovery in plants (Park et al., 2002; Reinhart and Bartel, 2002), miRNAs have been shown to play pivotal roles in a wide variety of metabolic and biological processes like plant growth, development, and response to biotic and abiotic stress (Axtell et al., 2007; Yang et al., 2007; Lu and Huang, 2008; Chuck et al., 2009; Lelandais-Briere et al., 2010).

In plants, miRNA genes are transcribed by a RNA polymerase II into a primary miRNA (pri-miRNA) with a cap and a poly (A) tail containing typical stem-loop regions and thereafter processed into a precursor miRNA (pre-miRNA). These two steps are processed by the RNAse III enzyme Dicer-like1 (DCL1) and its double-strand RNA binding (DRB) partner HYL1 to yield a double-stranded miRNA intermediate (duplex miR:miR*) (Bartel, 2004; Mallory et al., 2008; Vazquez et al., 2008). Alternatively, recently-evolved miRNAs genes use DCL4 and other DCL proteins to produce miRNAs (Mallory et al., 2008; Vazquez et al., 2008). After their transference from the nucleus to the cytoplasm via the transporter protein HASTY, miRNAs will be incorporated into an RNA-induced silencing complex (RISC), containing a protein of the Argonaute (AGO) family, where they will serve as a guide for sequence-specific recognition of protein-coding mRNAs and mediate their cleavage and/or inhibit their translation (Voinnet, 2009).

Since soybean (Glycine max) is one of the most economically important agricultural crops around the world, and one of the main global sources of protein and oil, both for food and for livestock feed, the interest in clarify the gene regulation network is crescent in this species. As miRNAs play important regulatory roles in a wide variety of developmental and metabolic processes in plants, they are being the focus of several researches in soybean (Zhang et al., 2008; Wang et al., 2009; Joshi et al., 2010; Kulcheski et al., 2011; Song et al., 2011; Turner et al., 2012; Zeng et al., 2012; Fang et al., 2013; Wang et al., 2013; Yan et al., 2013; Yin et al., 2013; Barros-Carvalho et al., 2014; Goettel et al., 2014; Guo et al., 2014; Wong et al., 2014; Xu et al., 2014). Soybean

yield is settled by the number of pods produced per unit area and individual seed weight (Kokubun, 2011). The seed number depends upon the number of floral buds that initiate pods and reach maturity (Kokubun, 2011). Usually, soybean generates a plentiful floral buds amount, however a significant proportion of them abort during development. Indeed, rates of flower and pod abscission/abortion were estimated to reach 80% (Shibles et al., 1975; Kokubun, 2011). In this way, to identify intrinsic soybean floral whorls miRNAs can be helpful for the understanding of floral gene regulation and consequently the soybean floral physiology. Joshi et al. (2010) performed a study involving the high-throughput sequencing of four different soybean tissues, including flower, and observed the expression profile of some miRNAs. However the analysis was done with the entire flower and was not possible to detect specific miRNAs from the different floral organs.

To acquire a better understanding about how miRNAs can act in different floral tissues, we decide to sequence three different soybean floral whorls. We selected the two sexual whorls, carpel and stamens, and also added petals samples, as these structures are close related to both reproductive organs, as petals and stamens share the B-class of MADS box genes for their organ identity (Huang et al., 2014).

The small RNA libraries of soybean carpel, stamen and petal tissues were constructed and analyzed. Conserved and novel soybean's miRNAs were identified and their expression profiles were confirmed by RT-qPCR. We also predicted the targets of some miRNAs. The results presented in this work expand our knowledge of the diversity and specificity of soybean miRNAs and provide a basis for further understanding of the biological mechanisms that take place in soybean floral tissues.

2. Material and methods

2.1. Plant material

In order to identify conserved and new miRNAs in soybean (*G. max* (L.) Merrill) and to verify those that are related to floral tissues, a set of three libraries of small RNAs were constructed from carpels, stamens and petals samples. These samples were collected from the soybean cultivar 'Urano' grown at the experimental field of Universidade de Passo Fundo (UPF) (Passo Fundo, Brazil) during the R2–R3 soybean developmental stage.

2.2. RNA extraction and sequencing

Total RNA from carpels, stamens and petals was isolated using TRIzol (Invitrogen, CA, USA) and the RNA quality was evaluated by electrophoresis in 1.0% agarose gel. The amount of the RNA was checked using a Quibit fluorometer and Quant-iT RNA assay kit according to the manufacturer's instructions (Invitrogen, CA, USA). Approximately 10 μg of total RNA was sent to Fasteris Life Sciences SA (Plan-les-Ouates, Switzerland) for processing and sequencing using Solexa technology on the Illumina Genome Analyzer GAII. Briefly, the processing by Illumina consisted of the following successive steps: acrylamide gel purification of RNA bands corresponding to the size range 20–30 nucleotides (nt), ligation of 3′ and 5′ adapters to the RNA in two separate subsequent steps each followed by acrylamide gel purification, cDNA synthesis and a final step of PCR amplification to generate DNA colonies template library for Illumina sequencing. After removing the vector the sequences with read lengths ranging from 18 to 26 nt were used for further analyses.

2.3. Sequence analysis

Herein, we employed the UEA sRNA workbench analysis tools (http://srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi) (Stocks et al., 2012) for processing the data obtained from the three small RNAs libraries. First of all, we selected the filter tool in order to discard all sequences that present the follow features: sequences smaller than 18 nt and longer that 26 nt; low complexity (sequences containing less than three distinct nucleotides), fragments of transfer and ribosomal RNAs (t/rRNAs) and also the sequences that were not mapped in the soybean genome. After that, we selected the *miRCat* (*miRNA categorization*) tool, which predicts mature miRNAs and their precursors from a small RNA (sRNA) dataset and a genome. The algorithm uses the technique described by Moxon et al. (2008) for predicting miRNA precursor hairpin. Once the sequences were mapped in the soybean genome, the loci were analyzed in order to find likely miRNA candidates. The most abundant sRNA read within a locus was chosen as the likely miRNA, flanking sequences surrounding this sRNA were extracted from the genome using 400 nt, and each precursor candidate was then folded using RNAfold.

The parameters used by the miRCat tool for classifying miRNAs are: the number of consecutive mismatches between miRNA and miRNA* must be no more than 3; the number of paired nucleotides between the miRNA and the miRNA* must be at least 17 of the 25 nucleotides centered around the miRNA; the hairpin must be at least 75 nt in length; the percentage of paired bases in the hairpin must be at least 50% of base-pairs in the hairpin. Finally the hairpin with the lowest adjusted minimum free energy (AMFE) from the sequence windows was chosen as the precursor miRNA (pre-miRNA) candidate.

Additionally, the sRNAs sequences were mapped on the miRNAs precursors candidates using the SOAP program (Li et al., 2008), which returns information concerning the alignment position. We considered miRNA precursors those which accumulated one or more reads blocks in the same orientation (Langenberger et al., 2012). Then, a representative sequence of each block (the most abundant in terms of read count) in the precursor, with at least 10 reads, was taken as the mature miRNA.

All the mature sequences were blasted against miRBase Release 21 in order to classify them into: novel soybean miRNAs families, known soybean miRNAs and new members of conserved soybean miRNAs families, and also isoforms of known soybean miRNAs. Then, the miRNAs were named as follow: NFxx (novel miRNAs families in soybean), gma-MIRxxx (known miRNAs) and gma-MIRxxx_iso (isoforms of known miRNAs). We follow the criteria and conventions for miRNA identification and naming described by Ambros et al. (2003).

The raw abundance values for sequence reads in every library were normalized into corresponding transcripts per million (TPM) abundance values using: TPM abundance = (raw value/#sum_use) * n_base, were n_base is a million (1,000,000) and #sum_use is number of total filtered reads among 18-26 nt. We also performed a nucleotide frequency analysis for the mature miRNAs identified by WebLogo (Crooks et al., 2004).

2.4. MiRNA validation and expression analysis by RT-qPCR

Reverse transcription quantitative polymerase chain reaction (RT-qPCR) amplification was performed to validate and investigate the expression of twelve new miRNAs (NF02ab-3p and 5p, NF04-3p and 5p, NF05, NF06-5p, NF10-5p2, NF15, NF18, NF19ab, NF20, NF22-3p) and eleven conserved miRNAs (gma-MIR159a-3p, gma-MIR164aefghijk-3p, gma-MIR166cdefgino-3p, gma-MIR166acl-5p, gma-MIR169afgm, gma-MIR169e, gma-MIR172abh-5p and gma-MIR5762-3p, gma-MIR156cde-5p, gma-MIR396bck-5p and gma-MIR9749) across the three different whorls. The cDNA synthesis was carried out by multiplex technique (Chen et al., 2005) from approximately 1 μg of total RNA. Each reaction was primed with a pool of 0.5 μM 10 gene-specific stem-loop primers. Before transcription, RNA and primers were mixed with RNase-free water to a total volume of 10 μL and incubated at 70 °C for 5 min followed by ice-cooling. Then, 6 μL 5× RT-Buffer (Promega, Madison, WI, USA), 1 μL of 5 mM dNTP (Ludwig, Porto Alegre, RS, Brazil) and 1 μL MML-V RT Enzyme 200 U (Promega, Madison, WI, USA) were added for a final volume of 30 μL. The synthesis was performed at 42 °C for 30 min on

Table 1
Summary of the sRNA sequences amount detected from soybean floral tissues libraries.

	Carpel	Stamen	Petal	Total
Read counts				
Total	3,530,011	3,275,049	3,360,601	10,165,661
18-26 nt	2,976,367	1,704,711	2,728,961	7,410,039
Removed t/rRNAs	136,510	349,966	288,874	775,350
Mapped to the soybean genome*	2,466,470	1,130,917	2,076,323	5,673,710
Unique read counts				
18-26 nt	1,671,684	554,189	1,235,627	4,020,267
Removed t/rRNAs	11,976	13,644	17,359	25,959
Mapped to the soybean genome	1,391,577	436,773	1,006,501	3,294,577

a Veriti Thermal Cycler (Applied Biosystems, Foster City, CA, USA), and inactivation of the enzyme was completed at 85 °C for 5 min. All cDNA samples were 100-fold diluted with RNase-free water before being used as a template in RT-qPCR analysis. The stem-loop primer, used for miRNA cDNA synthesis, was designed according to Chen et al. (2005), consisting of 44 conserved and six variable nucleotides that were specific to the 3′ end of the miRNA sequence (5′ GTCGTATCCA GTGCAGGGTCCGAGGTATTCGCACTGGATACGACNNNNNN 3′). The forward primers were designed based on the mature miRNA sequences, and the reverse primer was the universal reverse primer for miRNA (Chen et al., 2005). All primers sequences are presented in the Supplementary file 1.

The RT-qPCR reactions were performed in the Bio-Rad CFX384 real-time PCR detection system (Bio-Rad, Hercules, CA, USA) using SYBR Green I (Invitrogen, Carlsbad, CA, USA) to detect double-stranded cDNA synthesis. Reactions were completed in a volume of 10 μL containing 5 μL of diluted cDNA (1:100), 1× SYBR Green I (Invitrogen, CA, USA), 0.025 mM dNTP, 1× PCR Buffer, 3 mM MgCl₂, 0.25 U Platinum Taq DNA Polymerase (Invitrogen, CA, USA) and 200 nM of each reverse and forward primer. Samples were analyzed in biological triplicate in a 384-well plate, and a no-template control was included. We used gma-MIR167ef and the novel NF06-3p as reference genes, once they were the most stable genes across the three different tissues according to GeNorm analysis (Supplementary file 2). The conditions were set as the following: an initial polymerase activation step for 5 min at 95 °C, 40 cycles for 15 s at 95 °C for denaturation, 10 s at 60 °C for annealing and 10 s at 72 °C for elongation. A melting curve analysis was programmed at the end of the PCR run over the range of 65 to 99 °C, and the temperature increased stepwise by 0.5 °C. The threshold and baselines were manually determined using the Bio-Rad CFX manager software. To calculate the relative expression of the miRNAs we used the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). A Kruskal-Wallis statistical test was performed to compare the differences in expression among the different samples. The means were considered significantly different when $P < 0.05$.

2.5. Prediction of miRNA targets

Target prediction for miRNAs is straightforward because it is assumed that most of them match their targets with almost perfect complementarity (Bartel, 2009; Voinnet, 2009). The putative target genes for novel and conserved miRNAs that were differentially expressed by RT-qPCR were searched by using the web-based computer psRNA Target Server (http://biocomp5.noble.org/psRNATarget/) (Zhang, 2005). Mature miRNA sequences were used as queries to search for potential target mRNAs in the *G. max* database (DFCI gene index release 15). The total scoring for an alignment was calculated based on the miRNA length, and the sequences were considered to be miRNA targets if the total score were less than 3.0 points (mismatch = 1 and G:U = 0.5). Results from these analyses were individually

inspected on the Phytozome, where the loci and protein annotation were obtained.

3. Results and discussion

3.1. Sequencing and annotation of small RNAs

In order to identify novel as well as conserved miRNAs in soybean floral tissues, we generated three small RNAs libraries using pooled RNA isolated from carpels, stamens and petals. The high-throughput sequencing performed on Solexa platform generated a total of 10,165,661 reads from those 3,530,011 from carpel, 3,275,049 from stamen and 3,360,601 from petal tissues (Table 1). The following step was to select the reads ranging from 18–26 nt and remove the reads that corresponded to tRNAs and rRNAs. From a total of 6,634,689 sequences, 5,673,710 of them mapped to the soybean genome, from those 3,294,577 were unique sequences. The stamen library was the one with highest number of tRNAs and rRNA fragments (Table 1). Concerning to the 960,979 sequences that didn't match on soybean genome, we observed (in a previous study) that several of these small RNAs unrelated to the soybean genome were originated from viruses, bacteria and different groups of fungi sRNA libraries (Molina et al., 2012). This study, with a focus in metatranscriptomic, proved that high-throughput sequencing of soybean sRNAs can be a good source for detecting different pathogenic, symbiotic and free-living organisms.

After processing all these data using the UEA sRNA workbench tools and also the SOAP program we detected 200 miRNAs sequences across the three small RNAs libraries. The mature miRNAs were mainly represented by sequences with 21 nt (48% of the unique reads) and 22 nt (24%) (Fig. 1A). These miRNAs were processed from both arms of their precursors, being 52% from the 3p and 48% from the 5p arm (Fig. 1B). As previously reported (Aryal et al., 2014), we observed a higher accumulation of purine rich strands (sequences with more purine residue than pyrimidine residue) than the pyrimidine-rich strands (sequences with more pyrimidine residue than purine residue) in the miRNAs sequences (Fig. 1C). This phenomenon was already observed in other plant species (Aryal et al., 2014). To determine whether the purine–pyrimidine bias is localized on a particular region or if it is spread across the miRNAs, we analyzed the nucleotides frequency at each position in the 21, 22 and 24 nt long sequences (Fig. 1C). Uracil was the most frequent nucleotide at the 5′ end, which is in agreement with previous results for soybean miRNAs sequences (Subramanian et al., 2008; Kulcheski et al., 2011).

All the mature sequences were blasted against miRBase Release 21 in order to classify them into: novel soybean miRNAs families (Table 2), known soybean miRNAs (Table 3), new members of conserved soybean miRNAs families (Table 4) and also isoforms of known soybean miRNAs (Table 5). Then, the miRNAs were named as follow: NFxx (novel miRNAs families in soybean), gma-MIRxxx (known miRNAs) and gma-MIRxxx_iso (isoforms of known miRNAs). The only
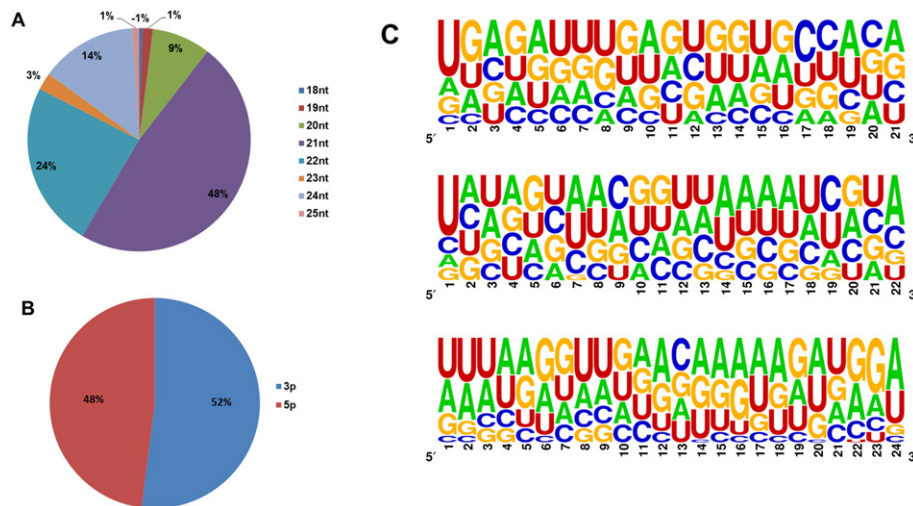
Fig. 1. The mature miRNA characterization. A) Sequence length, B) precursor's arm origin and C) position-specific nucleotide frequency among the most abundant mature miRNAs (21, 22 and 24 nt).

two sequences that were detected in our analysis but were not classified in these groups were the MIR5281 and gma-MIR1524-sibling (which will be further discussed). The MIR5281 is a miRNA already described in *Medicago truncatula* and *Brachypodium distachyon*, but was never observed in soybean before. This miRNA was only detected in carpel library; however we don't have information about their biological role in the other plant species.

We analyzed the presence of all 200 miRNAs across the three different libraries (Fig. 2). A total of 178 miRNAs were detected in all three libraries. Besides, four miRNAs (gma-MIR4355, NF04-3p, NF06-5p, NF32) were common to carpel and stamen libraries; eight miRNAs (gma-MIR4364a, gma-MIR4397-5p, gma-MIR5780-5p, gma-MIR394a-3p-iso, gma-MIR4411-iso, gma-MIR5036_iso, NF17, NF20) were common to carpel and petal, and five miRNAs (gma-MIR156di, gma-MIR169u, gma-MIR390b-3p, gma-MIR4345-3p, gma-MIR5762-5p) were common to stamen and petal tissues. Only one (NF19ab), two (NF-21, NF5281) and two miRNAs (gma-MIR398a-5p, gma-MIR828ab) were unique to stamen, carpel and petal libraries respectively.

Measures of RNA abundance are important for many areas of biology and often obtained from high-throughput RNA sequencing methods such as Solexa-Illumina sequence data. These measures need to be normalized to remove technical biases inherent in the sequencing approach, most notably the length of the RNA species and the sequencing depth of a sample (Wagner et al., 2012). To better analyze the miRNAs abundance, we convert the raw abundance values to normalized values by transcripts per million (TPM) abundance method. The frequency of miRNAs in each library will be discussed in the following sections according to the miRNAs classes.

3.2. Identification of novel soybean miRNAs families

Although the number of soybean miRNAs cataloged in the miRBase database release 21 has exceeded the number of 600 mature miRNAs, in this study we can prove that the discovery of new soybean miRNAs is no longer exceeded. A total of 41 novel miRNAs belonging to 32 novel families (Table 2) were identified by Solexa sequencing in libraries from carpel, stamen and petal of *G. max*. These families were provisory nominated as NF (new family) with a respective number (Table 2). Precursors of these novel miRNAs were identified, and they formed proper secondary hairpin structures, with MFEs ranging from −26.10 to −188.72 kcal/mol (Supplementary file 3). The most abundant mature miRNAs were 24 nt in length, which was interesting once that 21 nt length were the most common across all 200 detected sequences. We also evaluated the genomic location of the new miRNAs

(Supplementary file 4). Of the 41 new miRNAs genes identified in soybean, around 65% were located in intergenic regions and the rest were located inside of genes. From these, 32% were processed from only introns, while one new miRNA was processed from CDS-intronic junction region. More than 70% of the pre-miRNAs sequences were in the same sense direction (+) as the soybean genome annotation. For all 32 novel families identified, four (NF02ab, NF019ab, NF29ab, NF31ab) were compounded by miRNAs provided from two loci, and one (NF08abc) was processed from three different loci. For the rest, only one locus was detected. Sense and anti-sense miRNAs (miR/miR* or miR3p and 5p) were detected for eight families. As previously discussed by Kozomara and Griffiths-Jones (2011, 2014) the presence of miRNA and miRNA* to a given hairpin locus can provide robust and powerful discrimination between a bona fide miRNA and other transcribed fragment. According to the last release of miRBase, a growing body of evidence suggests that mature sequences derived from both arms of the hairpin may be biologically functional (Okamura et al., 2008), and even that the dominant mature sequence can be processed from opposite arms in different developmental stages, tissues or between orthologous miRNAs (Griffiths-Jones et al., 2011). The miRBase database recently ceased to use of the miR/miR* nomenclature, in favor of assigning names of the form xxx-5p and xxx-3p for sequences derived from the 5′ and 3′ arms of the hairpin precursor. So far, we decided to use the 3p or 5p nomenclature, instead miR and miR*, for all new identified miRNAs..

According to the abundance, we observed that the NF02ab-5p was the most abundant miRNA in petal tissues comparing to other novel miRNAs, while NF01-3p and NF04-5p were the most abundant in carpel and stamen respectively (Table 2).

3.3. Identification of conserved soybean miRNAs

In order to investigate the conserved soybean miRNAs (miRBase release 21) in carpel, stamen and petal tissues, we searched for known miRNAs in our small RNA libraries. A total of 639 *G. max* mature miRNAs are currently annotated at the miRBase. A total of 80 mature miRNAs were identified based on a perfect match between the sequenced reads and the annotated miRNAs deposited at the miRBase (Table 3), representing almost 13% of the currently known *G. max* miRNAs. For the families gma-MIR156, gma-MIR164, gma-miR167, gma-MIR169, gma-MIR171 e gma-MIR394 we identified new precursor's loci (Table 3). Most of the known mature miRNAs were 21 nt in length and processed from their 5p arm precursor. The gma-MIR166cdefgino-3p was the most abundant read in the carpel

Table 2
Novel soybean microRNA families identified in floral whorls by high-throughput sequencing.

Mature miRNA					Frequency TPM[c]		
miRNA ID[a]	Sequence (5′-3′)	Size (nt)	Precursor loci and strand[b]	miRNA read count	Carpel	Stamen	Petal
NF01-3p	AUAAGGCUUUGUUGUGGUAUUCCA	24	Gm14:558,487.558,711 [−]	101	30.41	2.65	11.08
NF02ab-5p	UUAUCAGUAGCAUCAUCAUCA	21	Gm07:10,001,901.10,002,013 [+]	106	3.65	14.15	39.01
			Gm07:10,004,513.10,004,625 [+]				
NF02ab-3p	GUUUGAUGAUGAUGUUACCGA	21	Gm07:10,001,901.10,002,013 [+]	45	1.22	10.61	14.45
			Gm07:10,004,513.10,004,625 [+]				
NF03-3p	UAACGACAGGUAGUAAAUUGAA	22	Gm15:17,904,646.17,904,916 [−]	77	16.22	6.19	14.45
NF04-5p	GUUUUUUUCUUUGUGGUGGCCGGG	24	Gm08:40,662,138.40,662,278 [+]	83	1.22	67.20	1.93
NF04-3p	CCGGCCACCAACAAAGAAAAAAACU	25	Gm08:40,662,138.40,662,278 [+]	13	0.41	10.61	0.00
NF05-3p	GUUUGGGGCUUGUGUUUUGUGGGC	24	Gm13:8,342,024.8,342,234 [+]	53	16.62	4.42	3.37
NF06-5p	AGGGUCACCUGUUCUUGGGUUA	22	Gm15:39,190,763.39,191,053 [−]	14	5.27	0.88	0.00
NF06-3p	ACCCAACAACAAGUGAUCCUUA	22	Gm15:39,190,763.39,191,053 [−]	49	15.00	3.54	3.85
NF07-3p	AACAUGGUAUCAGGGCCUGAUAGA	24	Gm16:19,925,619.19925903 [+]	12	3.65	1.77	0.48
NF08abc-3p	UAUUAAACGACCGAUGUAGAAAGU	24	Gm02:39,925,519.39,925,662 [−]	42	9.73	7.07	4.82
			Gm15:31,393,116.31,393,303 [−]				
			Gm18:42,937,102.42,937,265 [+]				
NF09-5p	UAGAUGAAGUUACUCUGAGCA	21	Gm14:6,304,120.6,304,369 [−]	17	2.03	1.77	4.82
NF09-3p	UACAUGUGCCUCUUCGUCGCUC	22	Gm14:6,304,120.6,304,369 [−]	53	9.33	16.80	5.30
NF10-5p1	AUUCGCACUGAAAUGGAUGUCCGU	24	Gm20:42,606,806.42,607,020 [+]	19	4.87	2.65	1.93
NF10-5p2	UUGUUGUGACUAGUUAAUGGGCAU	24	Gm20:42,606,806.42,607,020 [+]	38	10.14	5.31	3.37
NF10-3p	AUUAACUAGUCACAACAAUGGA	22	Gm20:42,606,806.42,607,020 [+]	14	4.05	0.88	1.44
NF11-5p	GCCGAAGAUGAAGAGCUUUGUAU	23	Gm16:7,236,797.7,236,894 [−]	37	13.38	2.65	0.48
NF12-5p	UUUGUUCUGGAUCCCUGUCGUC	22	Gm17:27,739,062.27,739,312 [+]	38	10.54	6.19	2.41
NF13-3p	UUGGCGGAAGUAAUACUAGGUA	22	Gm03:20,886,972.20,887,185 [−]	37	10.95	0.88	4.33
NF14-5p	AAGGCACCACUUCAGCAAUGGA	22	Gm14:4,584,761.4,585,058 [−]	23	5.27	0.88	4.33
NF14-3p	UUGCUGGACGUGGCGCCUUCCA	22	Gm14:4,584,761.4,585,058 [−]	12	2.03	0.88	2.89
NF15-5p	ACAGAUUGACAAUCCAUGUGAGCUA	25	Gm20:44,450,038.44,450,191 [−]	21	5.68	2.65	1.93
NF16-3p	AGUUAUCAAAAGCAAAAGUUUUGGA	24	Gm20:13,840,325.13,840,460 [+]	21	6.89	0.88	1.44
NF17-3p	GUUGAUGUGUCACAUGGAGAUGGA	24	Gm16:6,256,970. 6,257,129 [+]	23	2.03	0.00	8.67
NF18-3p	GAAUGGUGAGGAUGAAAAGUAACU	24	Gm10:37,397,199.37,397,385 [−]	23	5.68	0.88	3.85
NF19ab-5p	GUUGGACUCAAGGAACCUA	19	Gm07:44,231,008.44,231,088 [+]	8	0.00	7.07	0.00
			Gm17:493,663.493,743 [−]				
NF20-3p	UUAAGGACUAAAACAAAACAAACA	24	Gm04:5,821,402.5,821,618 [+]	10	3.65	0.00	0.48
NF21-5p	UGAGCCUUGCAGCAGUUAUGACAA	24	Gm10:47,656,000.47,656,302 [+]	8	3.24	0.00	0.00
NF22-5p	UGCUCAUUUUUAGUCCUGUAAGU	23	Gm08:11,432,998.11,433,200 [−]	8	2.43	0.88	0.48
NF22-3p	UUACAUGGACUAAAAAAUGAGCAAA	24	Gm08:11,432,998.11,433,200 [−]	52	14.60	2.65	6.26
NF23-3p	AAGGCAGAACGAUAUGUACGCAGA	24	Gm13:41,358,321.41,358,477 [+]	11	1.62	1.77	2.41
NF24-3p	UGAUUAUGAGGUCUGACACAAA	22	Gm13:11,972,958.11,973,052 [−]	45	6.89	7.96	9.15
NF25-5p	UCAUAGGAGAGAAAAAUAGGAAGG	24	Gm12:37,542,998.37,543,242 [−]	9	2.43	0.88	0.96
NF26-5p	UAUUGUGUAAGCUUCCUAAAGAGA	24	Gm01:8,573,823.8,573,906 [+]	16	0.81	0.88	6.26
NF27-3p	UAUUGGUCUUUUUGUAGUGAC	21	Gm02:6,233,520.6,233,676 [+]	19	0.41	2.65	7.22
NF28-5p	UAAAAUCGAUGUAGAAAGUGCC	22	Gm07:12,248,009.12,248,194 [+]	16	2.43	2.65	3.37
NF29b-5p	CCGUAUAACCAAACAUGGAGGGU	23	Gm19:40,112,991.40,113,140 [+]	16	3.24	4.42	1.44
NF29ab-3p	AUCUUGAUCGUUCAUGUUUUGGU	22	Gm14:25,154,100.25,154,318 [+]	16	2.43	0.88	4.33
			Gm19:40,112,991.40,113,140 [+]				
NF30-5p	GUAUUAGAUGCAUACUCAGUGGAU	24	Gm15:12,954,693.12,954,874 [+]	15	2.43	4.42	1.93
NF31ab-3p	UAUAGGACCGAUGUAGAAUGUGUU	24	Gm18:42,254,378.42,254,587 [−]	77	10.95	20.34	13.00
			Gm19:11,095,283.11,095,442 [−]				
NF32-3p	AUAACGAUUGUGCAAGUAUAGGGA	24	Gm16:1,163,892.1,164,088 [+]	7	2.03	1.77	0.00

[a] The number refers to a new family, the letter refers to the new member in this family and the 3p/5p refers to the mature sequence position into the arm precursor.

[b] (−) antisense and (+) sense strand.

[c] The normalization formula applied was: normalized frequency = miRNA count / total count of filtered reads × 1,000,000.

and stamen libraries. The gma-MIR159a-3p was the highest in reads accumulation from petal tissue, and also the second more abundant in carpel and stamen tissues. Some previous work already showed that both MIR159 and MIR166 are involved with floral organ development (Luo et al., 2013). In a previous work, Achard et al. (2004) proved that MIR159 is required for normal anther development by negative control of MYB33 and MYB65 (transcription factors involved in the gibberellin-promoted activation of LEAFY and in the regulation of anther development). In this study, the increasing of MIR159 expression caused a reduction in LEAFY transcript levels, with consequent flowering delay (in short-day photoperiods) and perturbation of anther development. This information corroborates with the high abundance of gma-MIR159a sequence in our stamen library. The targets of members of the MIR166 family are a Class III Hoemeodomain-leucine zipper (HD-ZIP III) transcriptional factors and their regulate an array of plant developmental processes, including shoot apical and lateral meristem formation, leaf polarity, floral and vascular development (Prigge et al., 2005).

Jung and Park (2007) observed that different genes of *Arabidopsis thaliana* MIR166 also exhibited differential expression domains in floral organs. Analyzing the MIR166 expression by the fusion of miRNA gene promoter with β-Glucuronidase (GUS), they observed that MIR166a was expressed predominantly in stamens, MIR166b was highly expressed in ovules and stigma, and MIR166d was expressed at high level in ovules. In contrast, MIR166g exhibited a broad spectrum of expression domains, being expressed in stigma, stamens, and receptacles but not in ovules. The diversity in the mature Ath-MIR166, where the sequences differ from one to a few single nucleotides, can be one of the reasons of a multiplicity in their physiological functions. We believe that this variability, also observed in the *G. max* MIR166 mature sequences, can be an explanation for the high abundance of gma-MIR166cdefgino-3p comparing to other members of the same family (like gma- MIR166acl-5p and gma-MIR166hk-3p).

Next, we analyzed any reads matching against the annotated precursor sequences. Surprisingly, in some cases, we found reads matching the

Table 3
Known soybean miRNA families detected in floral tissues by high-throughput sequencing.

Mature miRNA ID (miRBase)			New loci and strand[a]	Precusor arm	Read count[b]	Normalized frequency (reads/million)[c]		
miRNA ID	Sequence	Length (nt)				Carpel	Stamen	Petal
gma-MIR0156ah	UGACAGAAGAGAGUGAGCAC	20	Gm02:39,172,646.39,172,790 [+]	5p	612	41.76	127.33	175.79
gma-MIR0156b	UGACAGAAGAGAGAGAGCAC	20		3p	10	0.81	2.65	2.41
gma-MIR0156cde	UUGACAGAAGAUAGAGAGCAC	21		5p	1687	54.73	438.58	508.59
gma-MIR0156g	ACAGAAGAUAGAGAGCACAG	20		5p	11	1.22	2.65	2.41
gma-MIR0156pt	UUGACAGAAGAAAGGGAGCAC	21		5p	40	7.70	12.38	3.37
gma-MIR0159a-3p	UUUGGAUUGAAGGGAGCUCUA	21		3p	311,699	24,523.31	55,057.98	91,000.77
gma-MIR0159ae-5p	GAGCUCCUUGAAGUCCAAUU	20		5p	343	49.87	19.45	95.36
gma-MIR0159d	AGCUGCUUAGCUAUGGAUCCC	21		5p	52	10.54	7.07	8.67
gma-MIR0160af-5p	UGCCUGGCUCCCUGUAUGCCA	21		5p	362	85.95	84.00	26.49
gma-MIR0162a	UCGAUAAACCUCUGCAUCCA	20		3p	32	3.24	9.73	6.26
gma-MIR0162bc	UCGAUAAACCUCUGCAUCCAG	21		3p	1312	139.88	267.04	320.28
gma-MIR0164aefghijk	UGGAGAAGCAGGGCACGUGCA	21	Gm20:41,936,354.41,936,504 [−] Gm01:43,367,838.43,368,006 [+]	5p	1507	227.86	210.45	340.51
gma-MIR0166acl-5p	GGAAUGUUGUCUGGCUCGAGG	21		5p	3580	896.02	514.63	379.52
gma-MIR0166cdefgino-3p	UCGGACCAGGCUUCAUUCCCC	21		3p	166,652	24,858.20	52,525.52	22,124.69
gma-MIR0166hk-3p	UCUCGGACCAGGCUUCAUUCC	21		3p	12,553	1951.37	2505.93	2362.83
gma-MIR0167abd	UGAAGCUGCCAGCAUGAUCUA	21	Gm03:39,319,041.39,319,195 [+]	5p	2689	247.72	157.39	915.08
gma-MIR0167cj	UGAAGCUGCCAGCAUGAUCUG	21		5p	1301	10.95	81.35	569.28
gma-MIR0167ef	UGAAGCUGCCAGCAUGAUCUU	21		5p	5556	1609.99	1197.26	111.25
gma-MIR0167g	UGAAGCUGCCAGCAUGAUCUGA	22		5p	224	21.49	38.02	61.65
gma-MIR0167k	UGAAGCUGCCAGCCUGAUCUU	21		5p	218	0.41	8.84	99.70
gma-MIR0168a	UCGCUUGGUGCAGGUCGGGAA	21		5p	81,570	5316.91	16,283.25	24,100.78
gma-MIR0169afgm	CAGCCAAGGAUGACUUGCCGG	21		5p	39,208	1670.40	13,609.31	9486.48
gma-MIR0169b	CAGCCAAGGAUGACUUGCCGA	21		5p	82	10.14	38.02	6.74
gma-MIR0169c	AAGCCAAGGAUGACUUGCCGA	21	Gm15:14,202,424.14,202,597 [+]	5p	46	2.43	9.73	13.97
gma-MIR0169d	UGAGCCAAGGAUGACUUGCCGGU	23		5p	189	16.22	24.76	58.28
gma-MIR0169e	AGCCAAGGAUGACUUGCCGG	20	Gm09:5,299,542.5,299,763 [+] Gm13:371,080.371,201 [−] Gm15:14,188,469.14,188,641 [+] Gm15:14,191,146.14,191,325 [+] Gm17:4,864,158.4,864,292 [−]	5p	3525	134.20	984.16	1002.25
gma-MIR0169u	CAGCCAAGGAUGACUUGCCGU	21		5p	13	0.00	4.42	3.85
gma-MIR0171c	AGAUAUUGGUGCGGUUCAAUC	21		5p	8	1.22	2.65	0.96
gma-MIR0171efgju-3p	UGAUUGAGCCGUGCCAAUAUC	21		3p	1417	86.36	371.38	377.59
gma-MIR0171j-5p	UAUUGGCCUGGUUCACUCAGA	21		5p	2323	169.47	711.81	529.78
gma-MIR0171mt	UUGAGCCGCGUCAAUAUCUCA	21	Gm06:46,773,062.46,773,167 [−]	3p	7672	190.56	102.57	3412.76
gma-MIR0171np	UUGAGCCGCGUCAAUAUCUUA	21		3p	222	16.62	28.30	71.76
gma-MIR0172abh-3p	AGAAUCUUGAUGAUGCUGCAU	21		3p	213	10.95	15.92	80.91
gma-MIR0172cde	GGAAUCUUGAUGAUGCUGCAG	21		3p	163	8.51	118.49	3.85
gma-MIR0319do	UGGACUGAAGGGGAGCUCCUUC	22		3p	2008	426.52	273.23	311.61
gma-MIR0319f	UUGGACUGAAGGGGCCUCUU	20		3p	5591	1054.54	2273.38	201.80
gma-MIR0319hjkm	UUGGACUGAAGGGAGCUCCCU	21		3p	23,897	1585.67	9142.14	4646.19
gma-MIR0319n	UUUGGACCGAAGGGAGCCCCU	21		3p	469	181.23	11.50	4.33
gma-MIR0390afg-5p	AAGCUCAGGAGGGAUAGCGCC	21		5p	3744	268.40	1588.98	618.88
gma-MIR0390bd	AAGCUCAGGAGGGAUAGCACC	21		5p	151	4.05	110.53	7.71
gma-MIR0393hijk	UUCCAAAGGGAUCGCAUUGAU	21		5p	58	1.62	5.31	23.12
gma-MIR0394abcdefg-5p	UUGGCAUUCUGUCCACCUCC	20	Gm14:46,995,379.46,995,532 [−]	5p	488	30.81	88.42	150.27
gma-MIR0395defg	UGAAGUGUUUGGGGGAACUUU	21		3p	55	4.05	38.91	0.48
gma-MIR0396bck-5p	UUCCACAGCUUUCUUGAACUU	21		5p	11,621	242.86	2024.02	4205.99
gma-MIR0396bk-3p	GCUCAAGAAAGCUGUGGGAGA	21		3p	694	67.71	59.24	221.55
gma-MIR0396d	AAGAAAGCUGUGGGAGAAUAUGGC	24		3p	8	0.41	2.65	1.93
gma-MIR0396i-3p	GUUCAAUAAAGCUGUGGGAAG	21		3p	3431	45.41	580.94	1282.07
gma-MIR0396j-3p	AUUCAAGAUAGCUGUGGAAAA	21		3p	150	6.08	10.61	59.24
gma-MIR0398cd	UGUGUUCUCAGGUCGCCCCUG	21		3p	34	8.51	3.54	4.33
gma-MIR0399abch	UGCCAAAGGAGAGUUGCCCUG	21		3p	55	12.97	6.19	7.71
gma-MIR0399defg	UGCCAAAGGAGAUUUGCCCAG	21		3p	243	9.73	138.83	29.86
gma-MIR0482bd-3p	UCUUCCCUACACCUCCCAUACC	22		3p	2704	525.04	562.38	372.29
gma-MIR0530bcde	UGCAUUUGCACCUGCACUUUA	21		5p	17	2.03	5.31	2.89
gma-MIR0828ab	UCUUGCUCAAAUGAGUAUUCCA	22		5p	41	0.00	0.00	19.75
gma-MIR1507a	UCUCAUUCCAUACAUCGUCUGA	22		3p	79,954	5862.63	21,118.26	20,040.72
gma-MIR1508ac	UAGAAAGGGAAAUAGCAGUUG	21		3p	193	36.08	43.33	26.49
gma-MIR1510ab-3p	UGUUGUUUUACCUAUUCCACC	21		3p	539	30.41	176.85	127.15
gma-MIR1510b-5p	AGGGAUAGGUAAAACAACUACU	22		5p	2654	348.68	487.22	598.65
gma-MIR1513ab-5p	UGAGAGAAAGCCAUGACUUAC	21		5p	141	8.51	38.91	36.60
gma-MIR1515ab	UCAUUUUGCGUGCAAUGAUCUG	22		5p	77	3.65	9.73	27.45
gma-MIR2119	UCAAAGGGAGUUGUAGGGAA	21		3p	188	10.54	77.81	35.64
gma-MIR4355	CACUGUUGUGCUGGGGUGUACCA	22		3p	10	3.65	0.88	0.00
gma-MIR4358	CAGUGCAUGACUAUAUCGCCAG	22		3p	19	5.68	1.77	1.44
gma-MIR4364a	CGCGAGAUCGCACGGAAGAAGGUU	24		3p	15	1.62	0.00	5.30
gma-MIR4370	AGUAGACUCGUCCGAUUUUGCGUA	24		5p	10	2.43	0.88	1.44
gma-MIR4382	UAUGUUAACUGAUUUCAUGGAU	22		3p	92	9.73	17.68	23.12
gma-MIR4392	UCUGCGAAAAUGUGAUUUUCGGA	22		3p	186	34.06	72.51	9.63

Table 3 (continued)

Mature miRNA ID (miRBase)			New loci and strand[a]	Precursor arm	Read count[b]	Normalized frequency (reads/million)[c]		
miRNA ID	Sequence	Length (nt)				Carpel	Stamen	Petal
gma-MIR4397-5p	CAUCGUUGACGCUGACUGUACG	22		5p	14	1.62	0.00	4.82
gma-MIR4407	CAGAGGAAGCAGCACUUGUACC	22		3p	31	2.03	14.15	4.82
gma-MIR4408	UAACAACAUUGGAUGAGGGUUGGA	24		3p	39	10.95	6.19	2.41
gma-MIR4412-5p	UGUUGCGGGUAUCUUUGCCUC	21		5p	141	14.60	30.06	34.20
gma-MIR4414-5p	AGCUGCUGACUCGUUGGCUC	20		5p	8	2.03	1.77	0.48
gma-MIR4415ab-3p	UUGAUUCUCAUCACAACAUGG	21		3p	36	2.03	1.77	13.97
gma-miR5762-3p	TCATAGGAGGAATCAACTGGC	21		3p	43	0.00	15.03	12.52
gma-miR5773	TTTTTAAAAGGTTCAGTTAGGT	22		5p	21	6.08	0.88	2.41
gma-miR5780b-5p	TCTGAGTCCATGATATATTAAA	22		5p	10	3.24	0.00	0.96
gma-miR9722	TAATAGAGGGAAGAAGATGAA	21		5p	17	2.03	1.77	4.82
gma-miR9746abcdefgh	AAAGTGTTTGAATCTCAATTAGA	23		3p	17	0.41	10.61	1.93
gma-miR9749	TTAGCTTCTTTCACCTTTCCC	21		3p	33	2.43	16.80	3.85
gma-miR9757	CAACCCTCCTCAGTTAGATCTC	22		5p	16	5.27	0.88	0.96

[a] (−) antisense and (+) sense strand;
[b] Total reads count across the 3 libraries;
[c] The normalization formula applied was: normalized frequency = miRNA count / total count of filtered reads × 1,000,000.

miRBase-annotated precursor sequence, but reads matching the annotated mature miRNA sequences were absent. From this type of mature sequence we could observe that several of them were miRNAs antisense (or miRNA*) complementary to the soybean mature miRNAs present on miRBase. We list a total of 31 miRNAs (Table 4). Most of these sequences were 21 nt in length, and 50% were originated for each of 3p and 5p arms. The gma-MIR166g-5p was the most abundant miRNA in petal tissue, while gma-MIR166df-5p and gma-MIR156au-3p in carpel and stamen libraries respectively. The identification of these miRNAs contribute significantly for validating several of mature miRNAs that were already deposited on miRBase but so far have not had their complementary sequences identified.

Besides of miRNA*, the other sequences that were mapped in known precursors but haven't match in the mature miRNA were isoforms of miRNAs and also one miRNA-sibling. Isoforms of miRNAs and miRNA-siblings (Vazquez et al., 2008; Zhang et al., 2010) are miRNAs originated from the same primary transcript with distinct mature forms. MiRNA isoforms (isomiRs) often contain a small sequence shift or additional nucleotides (Ebhardt et al., 2010), whereas miRNA siblings are non-overlapping in comparison with the originally annotated mature or

Table 4
New miRNAs from floral tissues libraries, identified in the opposite arm of precursors previously deposited at miRBase.

miRNA ID	Sequence (5′-3′)	Length (nt)	New loci and strand[a]	Precursor arm	Total Read count[b]	Normalized frequency (reads/million)[c]		
						Carpel	Stamen	Petal
gma-MIR156au-3p_new	GCUCACUUCUCUAUCUGUCAGC	22	Gm02:39,172,646.39,172,790 [+]	3p	616	24.33	232.55	141.11
gma-MIR156di-3p_new	GCUCUCUAUACUUCUGUCAUC	21		3p	43	0.00	13.26	13.49
gma-MIR156k-3p_new	GCUCACUUCUCUUUCUGUCAAC	22		3p	9	1.62	1.77	1.44
gma-MIR164abdegk-3p_new	CACGUGCUCCCCUUCUCCAAC	21	Gm20:41,936,354.41,936,504 [-]	3p	14	3.24	1.77	1.93
gma-MIR166df-5p_new	GGAAUGGUGUCUGGUUCGAGA	21		5p	574	119.60	91.08	84.77
gma-MIR166e-5p_new	GGAAUGUUGGCUGGCUCGAGG	21		5p	377	90.41	62.78	39.97
gma-MIR166g-5p_new	GGAAUGUCGUUUGGUUCGAGA	21		5p	2782	8.51	164.47	1240.17
gma-MIR167bd-3p_new	GUCAUGCUGUGCUAGCCUCACU	22		3p	35	4.05	1.77	11.08
gma-MIR167c-3p_new	UCAGGUCAUCUUGCAGCUUCA	21		3p	10	0.81	1.77	2.89
gma-MIR167gj-3p_new	GAUCAUGUGGCUGCUUCACC	20		3p	71	0.41	12.38	26.97
gma-MIR168ab-3p_new	CCCGCCUUGCAUCAACUGAAU	21		3p	566	31.22	91.96	185.42
gma-MIR169a-3p_new	GGCAAGUUGUGUUUGGCUAUG	21		3p	165	1.62	64.55	42.38
gma-MIR169c-3p_new	GGCAGGUCAUCCUCUGGCUAUA	22		3p	24	1.62	3.54	7.71
gma-MIR171ai-5p_new	GGAUAUUGGUCCGGUUCAAUA	21		5p	58	14.19	3.54	9.15
gma-MIR319ab-5p_new	AGAGCUUUCUUCAGUCCACU	20		5p	33	2.03	9.73	8.19
gma-MIR398a-5p_new	GGAGUGAAUCUGAGAACACAAG	22		5p	7	0.00	0.00	3.37
gma-MIR398b-5p_new	GAGUGGAUCUGAGAACACAAGG	22		5p	362	52.30	75.16	71.28
gma-MIR1507a-5p_new	AGAGGUGUAUGGAGUGAGAGA	21		5p	186	31.22	38.02	31.79
gma-MIR1507b-5p_new	AGAGAUGUAUGGAGUGAGAGA	21		5p	219	42.57	32.72	37.08
gma-MIR1508ab-5p_new	ACUGCUAUUCCCAUUUCUAAA	21		5p	17	0.81	2.65	5.78
gma-MIR1511-5p_new	GUGGUAUCAGGUCCUGCUUCA	21		5p	410	70.95	116.72	49.61
gma-MIR4345-3p_new	CAAUCUUUUUAAGUUUCGUCU	21		3p	7	0.00	3.54	1.44
gma-MIR4349-3p_new	UUAUCUUUAGCCAAUGUGGGA	21		3p	90	12.57	21.22	16.86
gma-MIR4367-5p_new	ACUAGGGUUCAGGACAAUAUUCAA	24		5p	28	6.08	1.77	5.30
gma-MIR4395-3p_new	CAGCAGCUUUCUCGGACCCAUACU	24		3p	24	7.70	3.54	0.48
gma-MIR4407-5p_new	UAAAGUGUUGCUUCGUCUAAG	21		5p	7	0.41	4.42	0.48
gma-MIR5036-5p_new	CUCUCCCUCAAGGGCUUCUCG	21		5p	16	1.22	1.77	5.30
gma-MIR530acde-3p_new	AGGUGCAGGUGCAUCUGCAGG	21		3p	413	60.00	51.29	99.70
gma-miR5762-3p_new	CAGCUGAAUCCUCUUAUGAUC	21	Gm12:648,893.649,030 [+]	3p	12	0.00	3.54	3.85
gma-miR5780b-3p_new	UCAGGGAUUCAAAACAACGAAA	22	Gm13:34,948,270.34,948,416 [-]	3p	18	5.27	0.88	1.93
gma-MIR9735-3p_new	CCAAAGUUGGGCUUAAGCUGUA	22		3p	243	49.87	31.83	40.46

[a] (−) antisense and (+) sense strand;
[b] Total reads count across the 3 libraries;
[c] The normalization formula applied was: normalized frequency = miRNA count / total count of filtered reads × 1,000,000.

Table 5

miRNA isoforms detected in soybean by high-throughput sequencing.[a, b, c]

miRNA ID	Sequence	Length (nt)	New loci and strand[a]	Precursor arm	Total Read count[b]	Normalized frequency (reads/million)[c]		
						Carpel	Stamen	Petal
gma-MIR156bf_iso	UUGACAGAAGAGAGAGAGCAC	21		5p	249	21.08	103.46	38.53
gma-MIR159bcfzaa-3p_iso	AUUGGAGUGAAGGGAGCU	18		3p	16	1.22	0.88	5.78
gma-MIR166cdefgimno-3p_iso	GGACCAGGCUUCAUUCCCC	19		3p	154	26.35	50.40	15.41
gma-MIR167ef_iso	UGAAGCUGCCAGCAUGAUCUUA	22		5p	11,556	3735.30	1920.57	82.36
gma-MIR171ai_iso	UUGAGCCGUGCCAAUAUCACGA	22		3p	2448	296.37	206.03	714.73
gma-MIR171b-5p_iso	CGUGAUAUUGGUACGGCUCAUC	22		5p	202	5.27	15.92	82.36
gma-MIR172b-5p_iso	GUAGCAUCAUCAAGAUUCACA	21		5p	294	12.97	18.57	116.07
gma-MIR172cdeil_iso	GGAAUCUUGAUGAUGCUGCA	20		3p	61	0.41	51.29	0.96
gma-MIR319egmq_iso	UGUGCUUGGACUGAAGGGAGC	21		3p	51	10.14	13.26	5.30
gma-MIR319lp_iso	UUUGGACUGAAGGGAGCUCCU	21		3p	2735	682.76	823.23	57.79
gma-MIR390ac-3p_iso	CGCUAUCCAUCCUGAGUUUCA	21		3p	704	28.79	330.71	124.74
gma-MIR390b-3p_iso	CGCUAUCUAUCCUUGAGCUUCA	21		3p	18	0.00	12.38	1.93
gma-MIR393acdefghijk_iso	UCCAAAGGGAUCGCAUUGAU	20		5p	49	1.62	19.45	11.08
gma-MIR393c-3p_iso	GAUCAUGCUAUCCCUUUGGAU	21		3p	149	8.51	25.64	47.68
gma-MIR394a-3p_iso	AGCUCUGUUGGCUACACUUUG	21		3p	9	2.84	0.00	0.96
gma-MIR395abcdefgijklm_iso	UGAAGUGUUUGGGGGAACU	19	Gm01:4,788,545.4,792,485 [+] Gm08:40,838,512.40,838,659 [−]	3p	32	8.11	13.26	22.64
gma-MIR395abcijklm_iso	UGAAGUGUUUGGGGGAACUC	20	Gm02:1,772,395.1,772,531 [+]	3p	37	0.41	44.21	1.44
gma-MIR395abck_iso	UGAAGUGUUUGGGGGAACUCC	21	Gm02:1,730,678.1,730,791 [+]	3p	90	2.43	2.65	6.74
gma-MIR396abcehik-5p_iso	UUCCACAGCUUUCUUGAACU	20		5p	2210	45.00	624.27	670.90
gma-MIR398ab_iso	UUGUGUUCUCAGGUCACCCCU	21		3p	142	16.62	52.17	20.23
gma-MIR482ac-3p_iso	UUCCCAAUUCCGCCCAUUCCUA	22		3p	160	24.33	16.80	39.01
gma-MIR482bde_iso	UAUGGGGGAUUUGGGAAGGA	20		5p	818	160.96	126.45	133.89
gma-MIR482c-5p_iso	GGAAUGGGCUGAUUGGGAAG	20		5p	197	14.19	42.44	54.90
gma-MIR482c-5p_iso	GGAAUGGGCUGAUUGGGAAGU	21		5p	9867	859.53	1311.33	3016.87
gma-MIR1510a-3p_iso	UGUUGUUUUACCUAUUCCACCC	22		3p	1129	48.25	558.84	182.05
gma-MIR1510a-5p_iso	AGGGAUAGGUAAAACAAUGAC	21		5p	6219	478.82	1332.55	1700.60
gma-MIR1511_iso	AACCAGGCUCUGAUACCAUGG	21		3p	4687	562.75	717.12	1198.27
gma-MIR1512a-3p_iso	UGCUUUAAGAAUUUCAGUUAU	21		3p	11	1.62	1.77	2.41
gma-MIR1514a-5p_iso	UUCAUUUUUAAAAUAGGCAUUGGG	24		5p	19	2.43	2.65	4.82
gma-MIR1515ab_iso	UCAUUUUGCGUGCAAUGAUCU	21		5p	171	16.62	30.06	46.24
gma-MIR1516a-5p_iso	UUGGAUACAAGUUAUAAGCUCU	22		5p	52	12.97	6.19	6.26
gma-MIR1523a_iso	UAUGGGAUAAAUGUGAGCUCA	21		5p	151	8.51	21.22	51.05
gma-MIR1527_iso	UAACUCAACCUUACAAAACCGG	22		5p	29	9.33	0.88	2.41
gma-MIR2109-5p_iso	UGCGAGUGUCUUCGCCUCUGA	21		5p	2700	185.69	1036.33	515.33
gma-MIR4345_iso	CUAAGACGGAACUUACAAAGAU	22		5p	99	19.87	14.15	16.38
gma-MIR4349_iso	GUCCCAUAUUGGCUAGAGAUAAGA	24		5p	14	1.62	5.31	1.93
gma-MIR4351_iso	UUGGGAUUCAGUUGGAGUUGG	21		5p	36	7.70	4.42	5.78
gma-MIR4367_iso	UGAACCCUAGCGAAGUAAAUCA	22		3p	30	6.08	4.42	4.82
gma-MIR4376-5p_iso	UACGCAGGAGAGAUGACGCUG	21		5p	12	2.84	2.65	0.96
gma-MIR4395_iso	UAUGGGCUUGAGUAAGCUGCU	21		5p	8	1.62	0.88	1.44
gma-MIR4397-3p_iso	UCCCGUCAGUGUCAAAGAUGUG	22		3p	32	5.68	6.19	5.30
gma-MIR4411_iso	UAUUGUAACUAAUUUGUCGGUA	22		3p	10	2.43	0.00	1.93
gma-MIR4413a_iso	UAAGAGAAUUGUAAGUCACUG	21		5p	122	7.30	15.92	41.42
gma-MIR4416a_iso	UACGGGUCGCUCUCACCUAGG	21		3p	81	11.76	12.38	18.30
gma-MIR5036_iso	AGAGGCCCUUGGGGAGGAGUA	21		3p	9	0.81	0.00	3.37
gma-MIR5380ab_iso	UUAAGAAAAUGAAUGAUGAGGA	22	Gm15:18,047,158.18,047,244 [+] Gm16:11,320,145.11,320,309 [−]	3p	33	8.51	2.65	4.33

[a] (−) antisense and (+) sense strand.

[b] Total reads count across the 3 libraries.

[c] The normalization formula applied was: normalized frequency = miRNA count / total count of filtered reads × 1,000,000.

star miRNAs (Zhang et al., 2010). We detected 46 isoforms of miRNAs whose mature sequence is already known. We identified isoforms ranging from 18 to 24 nt. Interesting, some isoforms (with a high number of reads count) like gma-MIR482c-5p_iso and gma-MIR1511_iso had their mature miRNA missed in our sequencing, this fact was already related in the sequencing of miRNAs in the other legume *M. truncatula* (Devers et al., 2011) and also in soybean (Song et al., 2011). This can be an indicative that the isoforms can be the functional miRNA in determined situations. For example, in our libraries, the isoform gma-MIR167ef_iso presented higher abundance (11,556 reads) compared to the former mature gma-MIR167ef (5556 reads). This can be one of the reasons of discrepancies in miRNA expression analysis. If we consider that the oligonucleotides are developed for a determined miRNA sequence, in cases where miRNA sequence population change (ie. their isoforms), this specific sequence can no longer be detected. According to the

miRNA-sibling, the unique sequence identified in our libraries was the gma-miR1524-3p and its abundance was very low (Supplementary file 5).

3.4. Expression profile of conserved and newly identified soybean miRNAs

Knowledge about the expression patterns of miRNAs might provide clues about their functions. To get an insight into possible soybean miRNAs related to floral whorl, we examined the expression patterns of conserved and novel soybean miRNAs in petal, carpel and stamen tissues.

Using RT-qPCR, we observed that from the eleven conserved miRNAs, a set of eight miRNAs presented the same behavior across the three different tissues, while the expression of gma-MIR156cde-5p, gma-MIR396bck-5p and gma-MIR9749 were variable comparing

the different whorls (Fig. 3A). The expression of gma-MIR156cde-5p was higher in stamens comparing to carpel tissue. Gma-MIR396bck-5p presented a higher level in petal comparing to carpel. On the other hand, the specific soybean gma-MIR9749 was equally up-regulated in stamen and carpel comparing to petal tissues. These results were very interesting, since MIR156 and MIR396 are widely conserved plant miRNAs. We will discuss about these miRNAs function in the Subsection 3.5 when we discuss about their targets.

The RT-qPCR was also efficient to validate twelve novel soybean miRNAs. From these new miRNAs analyzed, we observed a differential expression in NF15, NF18 and NF22-3p in the floral tissues (Fig. 3 B). The new NF15 was differentially expressed across the all whorls. It was highest expressed in stamen and down-regulated in carpel. The NF18 demonstrated similar expression in stamen and carpel tissues but was significantly higher in these whorls compared to petals. The contrasting profile was observed for NF22-3p, where the lowest level was detected in carpel comparing to the same levels observed in petal and stamen tissue. These finds are very interesting, once that these new miRNAs can really be involved in the morphological and physiological roles that differentiate the soybean whorls.

We believe that the little difference observed on miRNAs expression across the different whorls was due the fact that samples were at the same biological stage (totally differentiated) when the soybean flowers were collected. Usually, the miRNAs that have being characterized as involved with floral biology were evaluated at meristems, during initial floral developmental phases (Achard et al., 2004; Luo et al., 2013).

3.5. Target prediction for the differentially expressed soybean miRNAs

In general, targets of plant miRNAs have perfect or near-perfect complementary site, allowing their identification using in silico prediction methods (Rhoades et al., 2002). We use psRNA target tool to search for target genes of the miRNAs that were differentially expressed by the RT-qPCR analysis (Supplementary file 6).

For the known MIR156, 26 loci were predicted as targets, from those 22 are codifying for the Squamosa promoter binding protein-like (SPB or SPL). The MIR156 family is one of the most evolutionary conserved miRNAs in plants as well their target proteins. In soybean, this miRNA target was already validated by two different approaches, degradome (Song et al., 2011) and modified 5′ RACE

assay (Shamimuzzaman and Vodkin, 2012) supporting the reliability of our computational analysis. This miRNA has already been extensively studied, and its biological function was well documented in *A. thaliana*. Ath-MIR156 also targets the transcriptional factors family SPL (Rhoades et al., 2002). In Arabidopsis was observed that MIR156 has its expression temporally regulated. Mature MIR156 is highly abundant in seedlings and decreases with time (Wu and Poethig, 2006). This expression pattern is observed not only in *A. thaliana*, but also in other species including *Arabis alpina*, *Cardamine flexuosa*, maize, poplar, rice, and tomato (Wang, 2014). The importance of MIR156 in flowering is inferred from the observation that overexpression of this miRNA delays flowering (Zhou and Wang, 2013). A differential expression of MIR156 across different floral organs was not mentioned before; however we observed a differential profile between stamen and carpel for this miRNA. For the other four predicted target loci, no functional annotation is provided.

The other conserved miRNA that was differentially expressed in the soybean floral whorls was gma-MIR396bck-5p. For this miRNA we predicted 35 targets codifying loci, from those 22 for Growth-regulating factors (GRF) proteins, one for each of the following proteins: DNA glycosylase, cysteine protease, glycosyl transferase, oxygenase, phosphatase and elongation factor proteins. Two loci were responsible for zein-binding and heat shock proteins. Although, we detected several different proteins, the MIR396 is a well-known regulator of the (GRF) proteins, which are responsible to promote and/or maintain cell proliferation activity in leaves. The GRF regulation by MIR396 was already
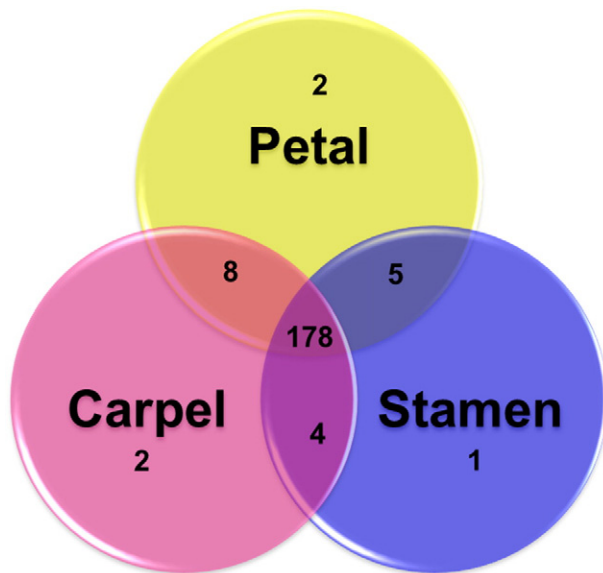

Fig. 2. Venn diagram representing the distribution of all 200 miRNAs identified in this study across the three different libraries.
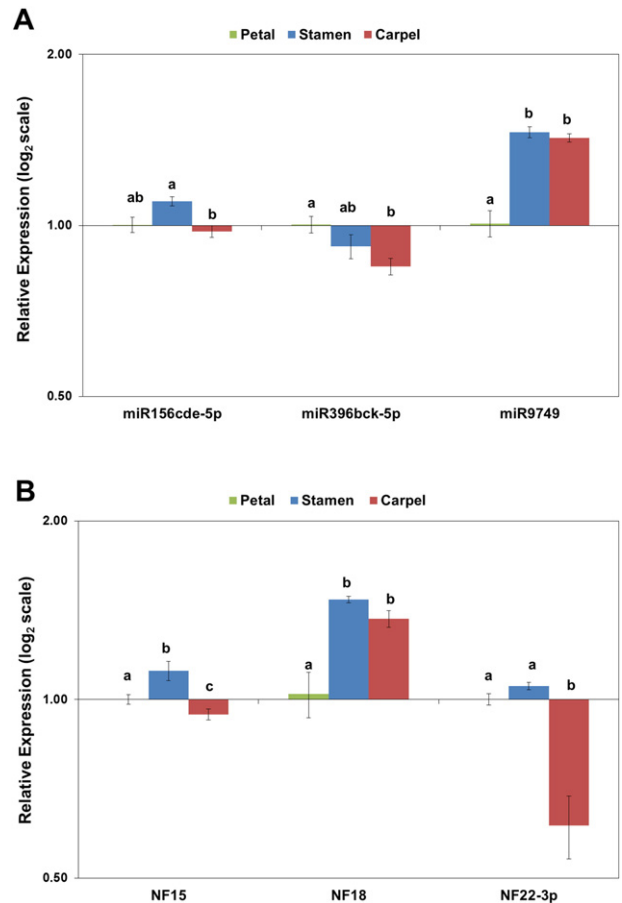

Fig. 3. Tissue-specific expression profiling of miRNAs identified in soybean whorls libraries by RT-qPCR. Expression profile of conserved (A) and newly (B) soybean miRNAs are represented by green, blue and red bars for petal, stamen and carpel tissues respectively. Error bars represent standard error of three independent biological replicates. The same letters "a" or "b" indicates no significant differences among the different tissues (Kruskal–Wallis test, $P < 0.05$).

demonstrated in soybean by degradome analysis (Song et al., 2011). This miRNA attenuates cell proliferation activity by repressing the targeted GRF genes (Wang et al., 2010). However in a recent study, MIR396 was also demonstrated to be involved in pistil development in Arabidopsis (Liang et al., 2014). The authors showed that the GRF proteins are also highly expressed in developing pistils and their expression are negatively controlled by the same miRNA that works in leaves tissues, i.e., MIR396. We observe that gma-MIR396bck-5p was higher expressed in petals than carpel tissues. This can be due the relationship between petal and leaves, once that these tissues share common evolutionary origins (Price et al., 2008).

For the species-specific gma-MIR9749 we detected 24 target loci responsible for Pentatricopeptide domain proteins. This miRNA was recently discovered in soybean cotyledons sRNA libraries (Goettel et al., 2014) and now in our libraries. Little is known about its biological function. We predicted its target and found several proteins belonging to a Pentatricopeptide repeat (PPR) proteins family. PPR proteins constitute one of the largest protein families in land plants, with more than 400 members in most species (Barkan and Small, 2014). Recent research has revealed that these proteins are involved with RNA editing and other RNA processing events in plant mitochondria and chloroplasts. PPR protein is a sequence-specific RNA-binding protein that identifies specific C residues for editing (Yagi et al., 2013). This indicates that this miRNA can be a good candidate for studies involving RNA-editing network.

We detected one predicted target loci for each of the two novel soybean miRNAs, NF15 and NF18 which were Glyma01g03020.1 and Glyma16g05860.1 respectively, however no biological function for both targets transcripts are available at Phytozome database. For the new NF15 target loci, we would like to point the interest in future approaches once that this miRNA was differentially expressed across the three different floral whorls. The new NF22-3p predicted target loci (Glyma08g02510.1 and Glyma05g37050.1) are codifying for E3 ubiquitin ligase proteins. This protein carry out the final step in the ubiquitination cascade, catalyzing the transfer of ubiquitin from an E2 enzyme to form a covalent bond with a substrate lysine (Berndsen and Wolberger, 2014). Ubiquitination by E3 ligases regulates several processes such as cell trafficking, DNA repair and signaling, being a key point in studies of cell biology.

4. Conclusions

This is the first study to systematically identify and annotate miRNAs in the different soybean floral whorls. We have identified a total of 200 mature miRNAs, from those 41 were novel miRNAs belonging to 32 new miRNAs families. The increased accumulation of specific and novel miRNAs, in the different whorls, indicates that these sRNAs are important components of the regulatory network in soybean flower. We also observed a differential gene expression of new and conserved miRNAs. The conserved miRNAs MIR156 and MIR396 have well-known targets with functions related to the reproductive process in plants. This study provides a new miRNA source for future research involving soybean flower tissues. The uncovering of gene regulation based on miRNAs network can contributes to understanding the molecular mechanisms involved in reproduction of soybean.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gene.2015.08.061.

Acknowledgments

References

Achard, P., Herr, A., Baulcombe, D.C., Harberd, N.P., 2004. Modulation of floral development by a gibberellin-regulated microRNA. Development 131, 3357–3365.
Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshal, M., Matzke, M., Ruvkun, G., Tuschl, T., 2003. A uniform system for microRNA annotation. RNA 9, 277–279.
Aryal, R., Jagadeeswaran, G., Zheng, Y., Yu, Q., Sunkar, R., Ming, R., 2014. Sex specific expression and distribution of small RNAs in papaya. BMC Genomics 15, 20.
Axtell, M.J., Snyder, J.A., Bartel, D.P., 2007. Common functions for diverse small RNAs of land plants. Plant Cell 19, 1750–1769.
Barkan, A., Small, I., 2014. Pentatricopeptide repeat proteins in plants. Annu. Rev. Plant Biol. 65, 415–442.
Barros-Carvalho, G.A., Paschoal, A.R., Marcelino-Guimaraes, F.C., Hungria, M., 2014. Prediction of potential novel microRNAs in soybean when in symbiosis. Genet. Mol. Res. 13, 8519–8529.
Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116, 281–297.
Bartel, D.P., 2009. MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233.
Berndsen, C.E., Wolberger, C., 2014. New insights into ubiquitin E3 ligase mechanism. Nat. Struct. Mol. Biol. 21, 301–307.
Chen, C., Ridzon, D.A., Broomer, A.J., Zhou, Z., Lee, D.H., Nguyen, J.T., Barbisin, M., Xu, N.L., Mahuvakar, V.R., Andersen, M.R., Lao, K.Q., Livak, K.J., Guegler, K.J., 2005. Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res. 33, e179.
Chuck, G., Candela, H., Hake, S., 2009. Big impacts by small RNAs in plant development. Curr. Opin. Plant Biol. 12, 81–86.
Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.
Devers, E.A., Branscheid, A., May, P., Krajinski, F., 2011. Stars and symbiosis: microRNA- and microRNA*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis. Plant Physiol. 156, 1990–2010.
Ebhardt, H.A., Fedynak, A., Fahlman, R.P., 2010. Naturally occurring variations in sequence length creates microRNA isoforms that differ in argonaute effector complex specificity. Silence 1.
Fang, X., Zhao, Y., Ma, Q., Huang, Y., Wang, P., Zhang, J., Nian, H., Yang, C., 2013. Identification and comparative analysis of cadmium tolerance-associated miRNAs and their targets in two soybean genotypes. PLoS One 8, e81471.
Goettel, W., Liu, Z., Xia, J., Zhang, W., Zhao, P.X., An, Y.Q., 2014. Systems and evolutionary characterization of microRNAs and their underlying regulatory networks in soybean cotyledons. PLoS One 9, e86153.
Griffiths-Jones, S., Hui, J.H., Marco, A., Ronshaugen, M., 2011. MicroRNA evolution by arm switching. EMBO Rep. 12, 172–177.
Guo, N., Ye, W., Yan, Q., Huang, J., Wu, Y., Shen, D., Gai, J., Dou, D., Xing, H., 2014. Computational identification of novel microRNAs and targets in Glycine max. Mol. Biol. Rep. 41, 4965–4975.
Huang, F., Xu, G., Chi, Y., Liu, H., Xue, Q., Zhao, T., Gai, J., Yu, D., 2014. A soybean MADS-box protein modulates floral organ numbers, petal identity and sterility. BMC Plant Biol. 14, 89.
Joshi, T., Yan, Z., Libault, M., Jeong, D.H., Park, S., Green, P.J., Sherrier, D.J., Farmer, A., May, G., Meyers, B.C., Xu, D., Stacey, G., 2010. Prediction of novel miRNAs and associated target genes in Glycine max. BMC Bioinf. 11 (Suppl. 1), S14.
Jung, J.H., Park, C.M., 2007. MIR166/165 genes exhibit dynamic expression patterns in regulating shoot apical meristem and floral development in Arabidopsis. Planta 225, 1327–1338.
Kokubun, M., 2011. Physiological Mechanisms Regulating Flower Abortion in Soybean, Soybean — Biochemistry, Chemistry and Physiology. In: Ng, P.T.-B. (Ed.), InTech.
Kozomara, A., Griffiths-Jones, S., 2011. MiRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. 39, D152–D157.
Kozomara, A., Griffiths-Jones, S., 2014. MiRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 42, D68–D73.
Kulcheski, F.R., de Oliveira, L.F., Molina, L.G., Almerao, M.P., Rodrigues, F.A., Marcolino, J., Barbosa, J.F., Stolf-Moreira, R., Nepomuceno, A.L., Marcelino-Guimaraes, F.C., Abdelnoor, R.V., Nascimento, L.C., Carazzolle, M.F., Pereira, G.A., Margis, R., 2011. Identification of novel soybean microRNAs involved in abiotic and biotic stresses. BMC Genomics 12, 307.
Langenberger, D., Pundhir, S., Ekstrom, C.T., Stadler, P.F., Hoffmann, S., Gorodkin, J., 2012. deepBlockAlign: a tool for aligning RNA-seq profiles of read block patterns. Bioinformatics 28, 17–24.
Lelandais-Briere, C., Sorin, C., Declerck, M., Benslimane, A., Crespi, M., Hartmann, C., 2010. Small RNA diversity in plants and its impact in development. Curr. Genomics 11, 14–23.
Li, R., Li, Y., Kristiansen, K., Wang, J., 2008. SOAP: short oligonucleotide alignment program. Bioinformatics 24, 713–714.
Liang, G., He, H., Li, Y., Wang, F., Yu, D., 2014. Molecular mechanism of microRNA396 mediating pistil development in Arabidopsis. Plant Physiol. 164, 249–258.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(−delta delta C(T)) method. Methods 25, 402–408.
Lu, X.Y., Huang, X.L., 2008. Plant miRNAs and abiotic stress responses. Biochem. Biophys. Res. Commun. 368, 458–462.
Luo, Y., Guo, Z., Li, L., 2013. Evolutionary conservation of microRNA regulatory programs in plant flower development. Dev. Biol. 380, 133–144.
Mallory, A.C., Elmayan, T., Vaucheret, H., 2008. MicroRNA maturation and action—the expanding roles of ARGONAUTEs. Curr. Opin. Plant Biol. 11, 560–566.
Molina, L.G., da Fonseca, G.C., de Morais, G.L., de Oliveira, L.F., de Carvalho, J.B., Kulcheski, F.R., Margis, R., 2012. Metatranscriptomic analysis of small RNAs present in soybean deep sequencing libraries. Genet. Mol. Biol. 35, 292–303.

Moxon, S., Schwach, F., Dalmay, T., Maclean, D., Studholme, D.J., Moulton, V., 2008. A toolkit for analysing large-scale plant small RNA datasets. Bioinformatics 24, 2252–2253.

Okamura, K., Phillips, M.D., Tyler, D.M., Duan, H., Chou, Y.T., Lai, E.C., 2008. The regulatory activity of microRNA* species has substantial influence on microRNA and 3′ UTR evolution. Nat. Struct. Mol. Biol. 15, 354–363.

Park, W., Li, J., Song, R., Messing, J., Chen, X., 2002. CARPEL FACTORY, a dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. Curr. Biol. 12, 1484–1495.

Price, A.M., Aros Orellana, D.F., Salleh, F.M., Stevens, R., Acock, R., Buchanan-Wollaston, V., Stead, A.D., Rogers, H.J., 2008. A comparison of leaf and petal senescence in wallflower reveals common and distinct patterns of gene expression and physiology. Plant Physiol. 147, 1898–1912.

Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N., Clark, S.E., 2005. Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in Arabidopsis development. Plant Cell 17, 61–76.

Reinhart, B.J., Bartel, D.P., 2002. Small RNAs correspond to centromere heterochromatic repeats. Science 297, 1831.

Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B., Bartel, D.P., 2002. Prediction of plant microRNA targets. Cell 110, 513–520.

Shamimuzzaman, M., Vodkin, L., 2012. Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing. BMC Genomics 13, 310.

Shibles, R.M., Anderson, I.C., Gibson, A.H., 1975. Soybean. In: Evans, L.T. (Ed.), Crop Physiology — Some Case Histories. Cambridge University Press, London, pp. 151–189.

Song, Q.X., Liu, Y.F., Hu, X.Y., Zhang, W.K., Ma, B., Chen, S.Y., Zhang, J.S., 2011. Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing. BMC Plant Biol. 11, 5.

Stocks, M.B., Moxon, S., Mapleson, D., Woolfenden, H.C., Mohorianu, I., Folkes, L., Schwach, F., Dalmay, T., Moulton, V., 2012. The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets. Bioinformatics 28, 2059–2061.

Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W.B., Zhu, J.K., Yu, O., 2008. Novel and nodulation-regulated microRNAs in soybean roots. BMC Genomics 9, 160.

Turner, M., Yu, O., Subramanian, S., 2012. Genome organization and characteristics of soybean microRNAs. BMC Genomics 13, 169.

Vazquez, F., Blevins, T., Ailhas, J., Boller, T., Meins Jr., F., 2008. Evolution of Arabidopsis MIR genes generates novel microRNA classes. Nucleic Acids Res. 36, 6429–6438.

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

Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 131, 281–285.

Wang, J.W., 2014. Regulation of flowering time by the miR156-mediated age pathway. J. Exp. Bot. 65, 4723–4730.

Wang, Y., Li, P., Cao, X., Wang, X., Zhang, A., Li, X., 2009. Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. Biochem. Biophys. Res. Commun. 378, 799–803.

Wang, L., Gu, X., Xu, D., Wang, W., Wang, H., Zeng, M., Chang, Z., Huang, H., Cui, X., 2010. miR396-targeted AtGRF transcription factors are required for coordination of cell division and differentiation during leaf development in Arabidopsis. J. Exp. Bot. 62, 761–773.

Wang, Y., Zhang, C., Hao, Q., Sha, A., Zhou, R., Zhou, X., Yuan, L., 2013. Elucidation of miRNAs-mediated responses to low nitrogen stress by deep sequencing of two soybean genotypes. PLoS One 8, e67423.

Wong, J., Gao, L., Yang, Y., Zhai, J., Arikit, S., Yu, Y., Duan, S., Chan, V., Xiong, Q., Yan, J., Li, S., Liu, R., Wang, Y., Tang, G., Meyers, B.C., Chen, X., Ma, W., 2014. Roles of small RNAs in soybean defense against *Phytophthora sojae* infection. Plant J. 79, 928–940.

Wu, G., Poethig, R.S., 2006. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. Development 133, 3539–3547.

Xu, M., Li, Y., Zhang, Q., Xu, T., Qiu, L., Fan, Y., Wang, L., 2014. Novel MiRNA and PhasiRNA biogenesis networks in soybean roots from two sister lines that are resistant and susceptible to SCN race 4. PLoS One 9, e110051.

Yagi, Y., Tachikawa, M., Noguchi, H., Satoh, S., Obokata, J., Nakamura, T., 2013. Pentatricopeptide repeat proteins involved in plant organellar RNA editing. RNA Biol. 10, 1419–1425.

Yan, Z., Hossain, M.S., Wang, J., Valdes-Lopez, O., Liang, Y., Libault, M., Qiu, L., Stacey, G., 2013. miR172 regulates soybean nodulation. Mol. Plant-Microbe Interact. 26, 1371–1377.

Yang, T., Xue, L., An, L., 2007. Functional diversity of miRNA in plants. Plant Sci. 172, 423–432.

Yin, X., Wang, J., Cheng, H., Wang, X., Yu, D., 2013. Detection and evolutionary analysis of soybean miRNAs responsive to soybean mosaic virus. Planta 237, 1213–1225.

Zeng, Q.Y., Yang, C.Y., Ma, Q.B., Li, X.P., Dong, W.W., Nian, H., 2012. Identification of wild soybean miRNAs and their target genes responsive to aluminum stress. BMC Plant Biol. 12, 182.

Zhang, Y., 2005. MiRU: an automated plant miRNA target prediction server. Nucleic Acids Res. 33, W701–W704.

Zhang, B., Pan, X., Stellwag, E.J., 2008. Identification of soybean microRNAs and their targets. Planta 229, 161–182.

Zhang, W., Gao, S., Zhou, X., Xia, J., Chellappan, P., Zhang, X., Jin, H., 2010. Multiple distinct small RNAs originate from the same microRNA precursors. Genome Biol. 11, R81.

Zhou, C.M., Wang, J.W., 2013. Regulation of flowering time by microRNAs. J. Genet. Genomics 40, 211–215.