Gene 575 (2016) 213–223

Contents lists available at ScienceDirect

Gene

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

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

ARTICLE INFO

*Article history:*
Received 7 April 2015
Received in revised form 23 July 2015
Accepted 28 August 2015
Available online 2 September 2015

*Keywords:*
MicroRNA
Soybean
Flower
Carpel
Stamen
Petal

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.



## 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

*Abbreviation:* miRNA, microRNA; nt, nucleotide; RT-qPCR, reverse transcription quantitative polymerase chain reaction; TPM, transcripts per million.

* Corresponding author at: Universidade Federal do Rio Grande do Sul, Centro de Biotecnologia, Laboratório de Genomas e Populações de Plantas (LGPP), Av. Bento Gonçalves, 9500-Setor IV-Prédio 43431, Sala 213, P.O. Box 15005, CEP 91501-970 Porto Alegre, RS, Brazil.
*E-mail address:* rogerio.margis@ufrgs.br (R. Margis).

http://dx.doi.org/10.1016/j.gene.2015.08.061
0378-1119/

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