

OPEN

Identification of microRNAs and relative target genes in *Moringa oleifera* leaf and callus

Stefano Pirrò^{1,2,5}, Ivana Matic^{1,3,5}, Arianna Guidi^{4,5}, Letizia Zanella^{3,5}, Angelo Gismondi^{3,5}, Rosella Cicconi⁴, Roberta Bernardini⁴, Vittorio Colizzi^{1,3}, Antonella Canini³, Maurizio Mattei^{4,5} & Andrea Galgani^{1,4*}

MicroRNAs, a class of small, non-coding RNAs, play important roles in plant growth, development and stress response by negatively regulating gene expression. *Moringa oleifera* Lam. plant has many medical and nutritional uses; however, little attention has been dedicated to its potential for the bio production of active compounds. In this study, 431 conserved and 392 novel microRNA families were identified and 9 novel small RNA libraries constructed from leaf, and cold stress treated callus, using high-throughput sequencing technology. Based on the *M. oleifera* genome, the microRNA repertoire of the seed was re-evaluated. qRT-PCR analysis confirmed the expression pattern of 11 conserved microRNAs in all groups. MicroRNA159 was found to be the most abundant conserved microRNA in leaf and callus, while microRNA393 was most abundantly expressed in the seed. The majority of predicted microRNA target genes were transcriptional factors involved in plant reproduction, growth/development and abiotic/biotic stress response. In conclusion, this is the first comprehensive analysis of microRNAs in *M. oleifera* leaf and callus which represents an important addition to the existing *M. oleifera* seed microRNA database and allows for possible exploitation of plant microRNAs induced with abiotic stress, as a tool for bio-enrichment with pharmacologically important phytochemicals.

Moringa oleifera Lam. is an arboreal medicinal plant of Moringaceae family, popularly called the “miracle tree”, widely distributed in the subtropical belt¹. Leaves, fruits, flowers and seeds of this plant are used as highly nutritious foods in many countries, particularly in India and Africa². It has been reported that the leaves of *M. oleifera* constitute a source of β -carotene, proteins, vitamin C, calcium, potassium, and may act as a source of antioxidants. Several experiments have demonstrated the ability of *M. oleifera* leaves to protect the body and the cells from oxidative stress linked to cancer and degenerative diseases³. Furthermore, *M. oleifera* is rich in anti-inflammatory molecules, effective in immune protection, thanks to the presence of quercetin and caffeoylquinic acid^{1,4}. Authoritative studies have shown that *M. oleifera* contains 46 antioxidants and 36 anti-inflammatory molecules, in addition to the omega-3-6-9 fatty acids^{4,5}.

Recently, plant tissue and cell suspension cultures have been investigated by biotechnological methods in order to provide a promising bioproduction platform for desired natural products. Medicinal plant *in vitro* culture consists of free cells or small groups of cells obtained from medicinal plant callus cultured in a liquid medium. This suspension can produce a massive standardized yield of secondary metabolites for pharmaceutical use, independently from geographical, environmental and seasonal factors^{6–9}.

MicroRNAs (miRNAs) are small (21–24 nucleotides), non-coding RNAs that modulate gene expression in eukaryotes. Inside the nucleus, miRNAs are transcribed by RNA polymerase II as precursor RNAs, known as primary miRNAs (pri-miRNAs)¹⁰. Subsequently, they are processed by DICER-LIKE 1 (DCL1) protein to release the pre-miRNAs. DCL1 also carries out the subsequent cleavage of pre-miRNAs, upon which the miRNA/miRNA* duplex is released and subsequently methylated at the 3' terminus by Hua Enhancer 1 (HEN1). Further on, it is exported to the cytoplasm, by the plant exportin protein HASTY^{10,11}. In the cytoplasm, miRNA/miRNA* duplex is separated and the guide strand is loaded into the RNA-induced silencing complex (RISC) through binding with Argonaute (AGO) proteins. This fully assembled RISC, which includes additional proteins such as Heat Shock

¹Mir-Nat s.r.l., Rome, 00133, Italy. ²Bioinformatics Unit, Centre for Molecular Oncology, Barts Cancer Institute, Queen Mary University London, London, EC1M 6BQ, UK. ³Department of Biology, University of Rome Tor Vergata, Rome, Italy. ⁴CIMETA, University of Rome Tor Vergata, Rome, Italy. ⁵These authors contributed equally: Pirrò Stefano, Matic Ivana, Guidi Arianna, Zanella Letizia, Mattei Maurizio and Gismondi Angelo. *email: galgani@scienze.uniroma2.it

Protein 90 (Hsp90), binds to the target transcript through sequence complementarity with its mature miRNA strand, to further direct either mRNA cleavage or translational inhibition^{12,13}.

MiRNAs are well-known for their role in regulating various plants processes under biotic and abiotic stresses. Many studies have confirmed that abiotic stress conditions can induce an enormous production of miRNAs in plants^{14,15}. In recent years, high-throughput sequencing and computational approaches have been used for identifying a large number of stress-related miRNAs. For example, in a study of Zhang *et al.*¹⁶, high-throughput sequencing analyses revealed cold stress-induced upregulation of 31 miRNAs in tea (*Camellia sinensis* (L.) Kuntze) plants. Similarly, Cao *et al.*¹⁷ studied the chilling stress response in wild tomato (*Solanum habrochaites*) and found that 192 miRNAs showed increased expression. Recently, Yang *et al.* found a different regulation of 84 miRNAs under low temperature stress in *Solanum aculeatissimum*¹⁸, while Zhou M. and Tang W. have reported that over-expression of rice miRNA156 (osa-miR156) results in increased cell viability and growth rate under cold stress in *Arabidopsis*, pine, and rice¹⁹. In our previous study, conserved and novel miRNAs in *M. oleifera* seeds were identified using Illumina platform technologies²⁰. However, results produced by our bioinformatics pipeline could have been inaccurate because of the unavailability of genome data at the time. Furthermore, there is no report on *M. oleifera* miRNAs response to cold stress. This study aims to identify the differential expression of conserved and novel miRNAs and their target genes in *M. oleifera* leaf, callus and cold stress treated callus by high-throughput small RNA sequencing. In addition, the miRNA repertoire of *M. oleifera* seed was re-evaluated in view of the recently published *M. oleifera* genome assembly²¹.

Methods

Callus production and cold stress treatment. *M. oleifera* seeds were deprived of their external envelope and sterilized by 1 min immersion in 70% ethanol. They were subsequently washed with sterile distilled water and soaked for 10 min in a 2.5% sodium hypochlorite solution with the addition of one drop of Tween 20. After 3–4 washes with sterile distilled water, seeds were ready to be sown in magenta boxes filled with medium containing full strength Murashige and Skoog (MS) basal medium (pH 5.8)²², 3% sucrose and 0.6% agar, previously sterilized by autoclaving at 121 °C for 20 min. Seed cultures were maintained in a growth chamber in a 16/8 h light/dark cycle at 22 °C.

Leaves from one-month old *Moringa* plants were sliced in uniformed pieces of 0.5 cm² in size. These pieces were placed with the abaxial page in contact with the substrate in magenta boxes filled with medium described above, supplemented with 0.5 mg/L of 2,4-dichlorophenoxyacetic acid. The culture was maintained in dark, in a growth chamber at 22 °C. Cold stress was applied to one-month old callus culture through the cultivation at 4 °C for 7 days. After the stress treatments, callus was reduced to powder using liquid nitrogen, pestle and mortar, and stored at –80 °C for subsequent analysis.

Small RNA sequencing. Next generation sequencing experiments, comprising samples quality control, were performed by Genomix4life S.R.L. (Baronissi, Salerno, Italy). Indexed libraries were prepared from 1 µg of purified RNA with TruSeq SmallRNA Sample Prep Kit (Illumina, CA, USA) according to the manufacturer's instructions. Libraries were quantified using the Agilent 4200 TapeStation (Agilent Technologies, CA, USA) and pooled to ensure that each index-tagged sample was present in equimolar amounts, with final concentration of 2 nM each. The pooled samples were subject to cluster generation and sequencing using an Illumina NextSeq 500 System (Illumina, CA, USA) in a 1 × 75 single read format at a final concentration of 3 pmol.

Pre-processing of sequencing data. Generated raw sequence files underwent quality control analyses using FastQC²³. Adapter sequences were trimmed using Cutadapt²⁴ with default parameters. The preprocessing tools included in the miRDeep2²⁵ suite of analysis were used for converting files in FASTA format, discarding reads shorter than 18 nt and collapsing the duplicates.

Prediction of high-confidence and novel miRNAs. mirDeep2²⁵ was used to predict mature, star and precursor sequences for well-annotated and novel miRNAs. Additionally, the secondary structure, as well as a detailed identification report, is produced for each predicted miRNA. In the analysis, the *M. oleifera* genome assembled by Tian *et al.*²¹ and *Arabidopsis thaliana* (L.) Heynh. miRNAs precursors list were used as references.

Prediction of low-confidence miRNAs. Blast alignment²⁶ (blastn version 2.6.0) respect to all plant-related, mature miRNAs annotated in miRBase²⁷ (release 21), was performed with the following parameters: *-task blastn-short*, *-perc_identity* 80, *-word_size* 15 and *-evalue* 1e⁻³. Identified matches were named according to the most abundant miRNA isoform highlighted in the comparison.

Evaluation of the conservation rate for low-confidence miRNAs, across all plant species. In a wide evolutionary context, known and low-confidence miRNAs from *M. oleifera* were compared with the whole plant-specific miRBase²⁷ [including 4 Coniferophyta (*Cunninghamia lanceolata* (Lamb.) Hook., *Pinus taeda* L., *Pinus densata* Mast., *Picea abies* (L.) H.Karst.), 1 Bryophyta (*Physcomitrella patens* (Hedw.) Bruch & Schimp.) 1 Lycopodiophyta (*Selaginella moellendorffii* Hieron.) 1 ancestral Magnoliophyta (*Amborella trichopoda* Baill.), 50 Magnoliophyta dicotyledons (including species of Asteraceae, Brassicaceae, Caricaceae, Cucurbitaceae, Euphorbiaceae, Fabaceae, Lamiales, Linaceae, Malvaceae, Ranunculaceae, Rhizophoraceae, Rosaceae, Rutaceae, Salicaceae, Solanaceae, Vitaceae family) and 12 Magnoliophyta monocotyledons (*Aegilops tauschii* Coss., *Brachypodium distachyon* (L.) Beauv., *Elaeis guineensis* Jacq., *Festuca arundinacea* Schreb., *Hordeum vulgare* L., *Oryza sativa* L., *Saccharum officinarum* L., *Saccharum* sp. L., *Sorghum bicolor* L. Moench, *Triticum aestivum* L., *Triticum turgidum* subsp. durum (Desf.) Husn., *Zea mays* L.) in order to compute the conservation frequency.

The conservation rate for a selected miRNA (c_i) is mathematically defined as:

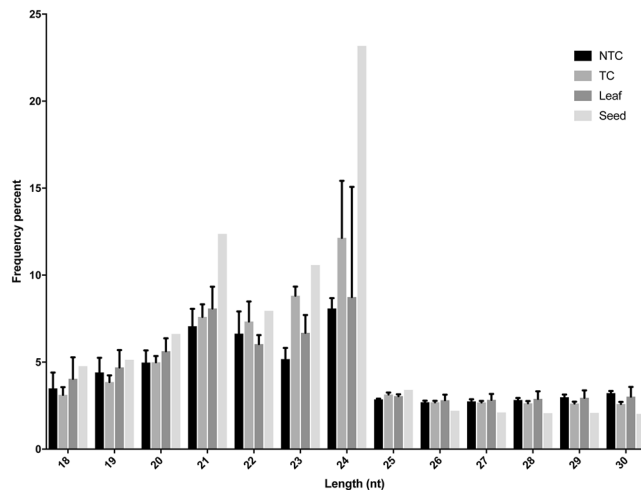


Figure 1. The length distribution of the clean and unique reads from leaf, seed and callus of *M.oleifera*. NTC: Non Treated Callus. TC: Treated Callus.

$$c_i = \sum_{n=0}^k [sequence_i = sequence_n]$$

where k = number of known plant-related miRNAs stored in miRBase²⁷

Prediction of miRNA targets. In order to predict the putative genes regulated by miRNAs identified in our samples, we took advantage of psRNATarget²⁸. The cDNA library used is the TAIR v10 of *A. thaliana*, while the scoring scheme follows the V2, released in 2017 by the authors.

Gene Ontology (GO) enrichment analysis. To better understand the biological functions underlying target genes, we performed enrichment analyses on the three categories of GO (biological processes, molecular functions, cellular compartments). GO annotation of the predicted targets was performed by uploading the list of highlighted genes in the PANTHER classification system²⁹.

Experimental validation of miRNAs using quantitative real-time PCR analysis. Total RNA was extracted from *M. oleifera* leaves, callus and cold stress treated callus using the NucleoSpin miRNA kit (Macherey-Nagel GmbH&Co., Germany), according to the manufacturer's protocol. Quality and quantity of the total RNA were evaluated by Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA) and by spectrophotometry (NanoDrop 2000, ThermoFischer Scientific, USA), respectively.

cDNA synthesis was performed with a miScript II RT Kit (Qiagen, USA) from 50 ng of sample. A quantitative real-time PCR analysis of mol-miR159a, mol-miR167a, mol-miR168a, mol-miR156a, mol-miR395a, mol-miR162a, mol-miR166i, mol-miR160h, mol-miR398b, mol-miR396a was performed using miScript SYBR Green PCR Kit according to the manufacturer's protocol (Qiagen, USA) and the reaction was performed using a Rotor-Gene Q (Qiagen, USA) machine. The amplification conditions were: activation/denaturation at 95 °C for 10 min followed by 40 cycles of denaturation at 94 °C for 10 sec, annealing at 60 °C for 30 sec and extension at 70 °C for 30 sec. All reactions were performed in triplicate for each sample and 5S rRNA was used as the internal control gene. Relative expression levels of miRNAs were quantified by using the $2^{-\Delta\Delta Ct}$ method as widely described in our previous work³⁰.

Results

Construction and sequencing of small RNA libraries. To identify *M. oleifera* miRNAs, total RNA was extracted from leaf, callus (Non Treated Callus; NTC) and cold stress treated callus (Treated Callus; TC) in order to construct nine small RNA libraries. Next, the libraries were sequenced using Illumina NextSeq. 500 System sequencing platform (Illumina, CA, USA) and analysed by bioinformatics approach. A total number of 71,900,401 raw reads was obtained from leaf (3 replicates): 72,962,350 from NTC (3 replicates) and 69,876,775 from TC (3 replicates). After removing the adaptors, low quality reads and contaminants, 63,698,781 clean reads from leaf, 71,445,523 from NTC and 68,889,789 from TC were obtained. As already reported in our previous work²⁰ the number of raw reads for seed tissue before and after cleaning process, was 31,290,964 and 22,737,895 - respectively. The clean and unique reads were subsequently subjected to size distribution analysis as shown in Fig. 1. The filtering analysis reported the length of clean and unique reads of small RNAs to be between 21 to 24 nucleotides (nt). The 24 nt was the most abundant size in seed and leaf (23% and 8.7%, respectively) followed by 21, 23 and 22 nt in size. In both NTC and TC, the most frequent small RNAs were 24 nt in length (8% and 12%, respectively), followed by small RNAs of 23, 21 and 22 nt in length, in TC and of 21, 22 and 23 nt in length in NTC.

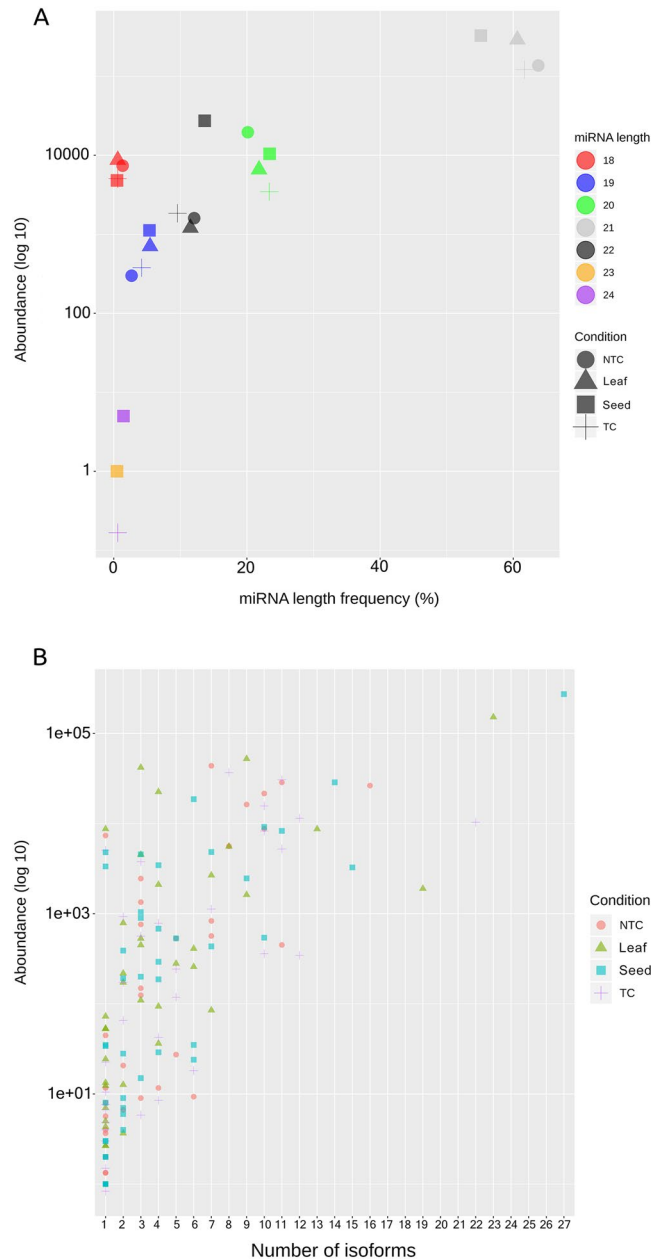


Figure 2. Size distribution vs abundance (A) and Number of isoforms vs abundance (B) in all the smallRNA libraries. The analysis was performed considering high and low confidence conserved smallRNAs in accordance for all experimental conditions. NTC: Non Treated Callus. TC: Treated Callus.

Known miRNAs in *M. oleifera*. As for the length distribution of mature miRNAs is concerned (high and low confidence miRNAs), 21 nt is the most frequent class with the majority reads in all samples, followed by 20, 22, 19 and 18 nt. The majority of miRNAs from leaf belongs to 21 nt category (61%, 2.9×10^5 total number of reads). The abundance of reads supporting this length is notably higher also in seed (55.1%) with a great number of reads (3.2×10^5). NTC and TC presented a highly frequent 21 nt miRNAs size (63.8% and 61.7%, respectively) with a high number of reads (1.3×10^5 and 1.2×10^5 , respectively) (Fig. 2A).

All the predictions obtained by the analysis with mirDeep2 software²⁵ were classified as high-confidence. Table 1 shows the most abundant high-confidence miRNAs (over 100 reads) in all experimental conditions. For each of them, a detailed set of sequence (mature, star, precursor) and structural information is reported in Supplementary Table 1. As for the expression of high-confidence miRNA isoforms, the most abundant in NTC is mol-miR398c_1 with 4.3×10^4 reads (mean value of triplicate). In TC, mol-miR398c_3 is the most frequent isoform with 3.1×10^4 reads (mean value of triplicate) while in leaf and seed, mol-miR166u_5 and mol-miR166u_2 with 1.6×10^5 and 1.3×10^5 reads (mean value of triplicate), respectively, are the most abundantly expressed isoforms (Supplementary Table 1).

miRNA family	miRNA name	Sequence	Read Counts			
			NTC	TC	LEAF	SEED
High confidence miRNAs						
156	156f ^{1,2,4}	TGACAGAAGAGAGTGAGCAC	399	693	589	2,427
	156f ⁵	TGACAGAAGAGAGTGAGCA	0	0	331	0
	156t ¹⁻³	TTGACAGAAGAAAAGAGAGCAC	0	0	99	126
157	157a ¹⁻⁸	TTGACAGAAGATAGAGAGCAC	1,797	674	64,679	1,044
160	160h ¹⁻⁵	TGCTGGCTCCCTGTATGCCA	43	26	446	0
162	162a ¹⁻⁴	TCGATAAACCTTGCATCCAG	206	284	351	176
166	166e ¹⁻⁵	TCGGACCAGGCTTCATTCCCC	8,026	10,774	55,612	65,352
	166i ¹⁻²	TCGGACCAGGCTTCATTCCCC	8,021	10,752	55,668	65,306
	166u ¹⁻⁵	TCTCGGACCAGGCTTCATTCC	25,234	24,900	202,663	137,652
167	167d ^{1-6,11}	TGAAGCTGCCAGCATGATCTGA	203	135	297	7,449
	167d ^{7-10,12,13}	TGAAGCTGCCAGCATGATCT	115	199	311	0
	167h ¹⁻³	TGAAGCTGCCAGCATGATCTT	141	124	1,060	18,461
171	171a ¹⁻⁹	TGATTGAGCCGTGCCAATATC	0	7	7,464	459
319	319 ¹	TTGGACTGAAGGGAGCTCCCT	0	7,082	0	0
	319 ^{2,3}	TTGGACTGAAGGGAGCTCCC	17,830	10,551	2,543	1,158
	319e ¹	TTGGACTGAAGGGAGCTCC	197	376	25	0
	319e ²	TTGGACTGAAGGGAGCTCCT	0	0	14	983
	319e ^{3,4}	TGGACTGAAGGGAGCTCC	215	0	0	0
	319g ¹	TGGACTGAAGGGAGCTCC	103	0	0	0
390	390a ¹⁻⁵	AAGCTCAGGAGGATAGCGCC	4,251	11,710	370	3,471
	390d ¹⁻⁴	AAGCTCAGGAGGATAGCGCC	1,336	6,668	114	1,171
393	393a ¹⁻²	TCCAAAGGGATCGCATTGATC	75	101	89	1,301
394	394b ¹⁻³	TTGGCATTCTGTCCACCTCC	21	60	2,102	3,444
395	395a ¹	GTCCCCCTGAGCACTCACC	0	0	0	257
396	396h ¹	TTCCACAGCTTTCTTGAACCTG	8,050	12,762	3,256	891
397	397a ^{1,2}	TCATTGAGTGCAGCGTTGATG	20,281	14,614	777	631
398	398c ¹⁻³	TGTGTTCTCAGGTCGCCCTG	43,595	88,821	22,492	3,343
399	399a ^{1,2,5-8,11,12}	CGCCAAAGGAGAGTTGCCCTT	90	195	130	0
	399a ^{3,4}	CGCCAAAGGAGAGTTGCCCT	0	0	0	290
	399a ^{9,10}	CGCCAAAGGAGAGTTGCC	0	3	130	0
	399d ^{5,7}	TGCCAAAGGAGAGTTGCCCTT	0	0	12	103
535	535a ¹⁻³	TGACAACGAGAGAGAGCACGC	65	310	853	389
	535a ⁴	GTGCTCTCTATCGTTGTCA	0	0	122	0
827	827 ¹	TTAGATGACCATCAACAAACG	735	1047	12	27
Low confidence miRNAs						
156	156 ¹	TTGACAGAAGATAGAGAGC	12	12	121	8
	156a ²	TGACAGAAGAGAGTGAGCAC	0	0	104	0
	156c ¹	TTGACAGAAGAAAAGAGAGCAC	22	18	0	473
	156e ¹	TTGACAGAAGAGAGTGAGCAC	117	139	350	87
	156g-3p ²	GCTCTCTATTCTTCTGTATC	0	0	122	0
	156z ¹	TGACAGAAGATAGAGAGCAC	7	4	755	14
159	159 ¹	TTGGATTGAAGGGAGCTCTA	27,752	31,510	51,769	9,022
	159b-3p ¹	TTGGATTGAAGGGAGCTCTT	108	143	471	63
	159c-3p ¹	CTTGGACTGAAGGGAGCTCCC	737	901	52	22
164	164a ¹	TGGAGAAGCAGGGCACGTGCA	7	9	7	180
166	166a ¹	TCGGACCAGGCTTCATTCC	84	54	106	666
	166a-5p ¹	GGAATGTTGTCTGGCTCGAGG	41	62	23	204
	166b ¹	TCGGACCAGGCTTCATTCCC	186	182	143	319
	166d ¹	TCGGACCAGGCTTCATTCCCC	58	105	445	554
	166e ¹	GGACCAGGCTTCATTCCCC	1	1	4	376
	166e-3p ¹	CTCGGACCAGGCTTCATTCCC	32	87	18	1,271
	166m ¹	CGGACCAGGCTTCATTCCCC	1	8	6	1,527
	166u ¹	TCTCGGACCAGGCTTCATTCC	24	19	149	172
Continued						

miRNA family	miRNA name	Sequence	Read Counts			
			NTC	TC	LEAF	SEED
167	167-3p ¹	AGATCATGTGGCAGTTTCACC	0	0	0	247
	167b ²	TGAAGCTGCCAGCATGATCTGG	0	2	277	33
	167b ³	TGAAGCTGCCAGCATGATCTGA	0	0	3	132
	167d ¹	TGAAGCTGCCAGCATGATCTTA	30	33	200	2,237
168	168 ¹	TCGCTTGGTGCAGGTCGGGA	34	83	21	139
	168a ¹	TCGCTTGGTGCAGGTCGGGAA	108	235	66	332
171	171-5p ¹	TATTGGCCTGGTTCACCTCAGA	1	1	463	64
319	319a ¹	TTGGACTGAAGGGAGCTCCCT	3,480	2,146	101	150
	319a ³	TTGGACTGAAGGGAGCTCCC	0	2,616	0	0
	319p ¹	TTTTGGACTGAAGGGAGCTCC	2	2	1	166
390	390a-3p ²	CGCTATCCATCCTGAGTTTCA	0	113	1	27
	390b-5p ¹	AGCTCAGGAGGGATAGCGCC	7	136	9	169
393	393-5p ¹	TTCCAAAGGGATCGCATTGAT	6	8	7	172
	393b ¹	TCCAAAGGGATCGCATTGATCT	324	674	204	9,745
	393b-5p ¹	TTCCAAAGGGATCGCATTGATC	421	705	150	7,445
396	396-5p ¹	TTCCACAGCTTCTTGAAGT	518	892	223	267
	396a-3p ¹	GTTCAATAAAGCTGTGGGAAG	37	124	28	2
	396b ¹	TTCCACAGCTTCTTGAAGTT	106	238	5,033	7,061
	396e-3p ¹	CTCAAGAAAGCTGTGGGAGA	2	4	113	41
397	397b ¹	TCATTGAGTGCAGCGTTGATGT	494	449	14	48
398	398f ¹	GGTGTTCACAGGTCGCCCTG	61	103	25	0
403	403 ¹	TTAGATTCACGCACAACTCGT	70	117	123	85
	403a ¹	TTAGATTCACGCACAACTCG	2,385	4,965	4,403	4,469
408	408 ⁴	ATGCACTGCCTCTCCCTGGC	0	2,278	0	171
	408d ¹	TGCACTGCCTCTCCCTGGC	465	944	198	27
530	530 ¹	TCTGCATTGCACCTGCACCT	4	9	441	893
6300	6300 ¹	GTCGTTGTAGTATAGTGG	7,374	2,741	8,719	4,809

Table 1. Summary of the most abundantly expressed miRNAs in different experimental conditions. Data are reported as the sum of identical mature sequence isoforms. NTC: Non Treated Callus. TC: Treated Callus.

A subsequent secondary comparison between the sRNA sequences of ten libraries and known mature miRNA sequences from other plants was performed. The outcome of this analysis generated a total of 300 low-confidence miRNAs that share their mature sequence with miRBase (Release 21)²⁷ records. All of them belong to a total of 56 miRNA families (Table 1, Supplementary Table 2). In most families, more than one precursor was identified. The most abundant isoforms resulted to be mol-miR159_1 in NTC, TC and leaf with 2.7×10^4 ; 3.1×10^4 and 5.1×10^4 reads (mean value of triplicate), respectively, and mol-miR393b_1 in seed with 9.7×10^3 reads (mean value of triplicate) (Supplementary Table 2).

In high and low confidence miRNA families, mol-miR166 showed a higher number of isoforms in seed (27 isoforms, 2.7×10^5 total number of reads), in leaf (23 isoforms, 1.5×10^5 total number of reads) and in TC (22 isoforms, 1.0×10^4 total number of reads), while in NTC, miRNA166 families possessed 16 isoforms with high abundance (2.6×10^4 total number of reads). Thirty-six percent of miRNA families has only one condition-specific isoform: among these, mol-miR6300 has the higher number of reads in leaf (8,719 reads), callus (7,374 and 5,057 reads in NTC and TC, respectively) and seed (4,809 reads) while mol-miR398 showed high expression only in seed (3,343 reads) (Fig. 2B).

As for low-confidence miRNAs, we assessed the correlation between conservation rate and abundance in all experimental conditions. mol-miR159_1 was the most abundantly expressed in *M. oleifera* leaf, NTC and TC (5.1×10^4 , 2.7×10^4 and 3.1×10^4 reads respectively) with a low conservation rate among plants (rate 30). Mol-miR156a_2 was found to be the most conserved miRNA among plants (rate 165), although uncommon in all experimental conditions except for the leaf (104 reads). In addition, mol-miR395a_2, mol-miR171a_1, mol-miR164a_1, mol-miR172a_1, mol-miR167a_2 and mol-miR169b/l_1 showed a high conservation rate among plants and a relatively low abundance, except for mol-miR164a_1 in the seed. The remaining miRNAs, with the exception for mol-miR396b_1, showed a conservation rate below 50 with different abundance degree. (Fig. 3).

Novel miRNAs in *M. oleifera*. A total of 392 novel microRNAs were predicted by identifying all the potential precursors (pre-miRNAs) and modelling their secondary structures. All the discovered sequences ranged in length between 18 and 25 nt (Supplementary Table 3). The sequences of most of the novel miRNAs were 21 nt long (102 sequences) while the second most abundant class was 18 nt long (69 seq). The lengths of novel miRNA precursors ranged from 37 to 105 nt, where the 43 nt was most abundant (17 seq). Leaf has the highest number of novel miRNAs (213 seq). NTC, TC and seed follow with 121, 85 and 56 sequences, respectively. The

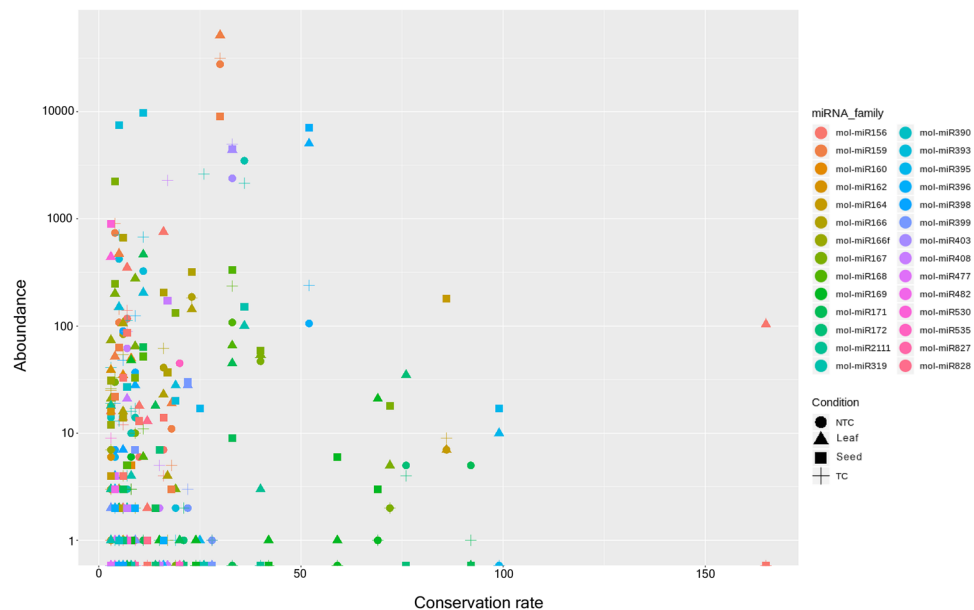


Figure 3. Conservation rate vs abundance of conserved miRNA families. The analysis was performed considering low confidence conserved top 100 miRNAs in accordance for all experimental conditions. NTC: Non Treated Callus. TC: Treated Callus.

most abundant miRNA isoform was found in the leaf with 21 nt in length (mol-miR-n31/n211/n239; 7.8×10^5 reads), followed by seed (mol-miR-n323; 2.0×10^5 reads), TC (mol-miR-n217; 1.8×10^4 reads) and NTC (mol-miR-n290; 1.4×10^5 reads) with 20 nt in length.

Identification of “cold-sensitive” miRNAs. Cold stress has been shown to have an important influence on plant growth and development by modulation of gene expression. Some studies document that chilling stress increases the amount/slow down decomposition of polyphenols in some plants (*Nicotiana tabacum* L., *Vitis vinifera* L.)^{31,32}. Moreover, chilling stress response induced increased expression of selected miRNAs in wild tomato (*Solanum habrochaites*)¹⁷, *Brassica rapa* L.³³ and *Astragalus membranaceus*³⁴. We performed a comparative analysis between the known miRNAs identified in *M. oleifera* and the ones published in the aforementioned studies. As shown in Fig. 4A, 25 miRNAs are conserved among the selected species. By looking at their global abundance within different conditions (Fig. 4B), it is not possible to determine whether the subset in its entirety is sensitive to the cold stress. However, miR-166e-3p in *Brassica rapa* L., mol-miR319_1 and mol-miR396h_1 in *M. oleifera*, ame-miR396-5 in *A. membranaceus* are upregulated after cold stress treatment.

Identification of condition-specific and backbone miRNAs. In addition to analysis of the gene regulation mediated by miRNAs, the analysis of differential internal regulation of miRNA expression in different tissues and experimental conditions was performed. In plants, this fine-tuning mechanism is based on the regulation of miRNAs decoy time³⁵. As consequence, miRNA patterns in plants are modulated according to the tissue and the environmental conditions. On the other hand, we also defined a set of miRNAs whose expression is stable across different experimental conditions (backbone). In order to evaluate the modulation of miRNA patterns among tissues and cold-stress, we overlaid the high-confidence, low-confidence and novel miRNAs identified in each condition, separately (Fig. 5). A detailed list of tissue-specific and backbone miRNAs is also reported in the Supplementary Table 4. Among 56 miRNA families belonging to all experimental conditions, mol-miR166 and mol-miR159 families are the largest families identified with ten and eight members, respectively. A total of five different isoforms belongs to the mol-miR156 and mol-miR396 families. Several miRNAs appear to be tissue-specific: 135 unique miRNAs were found in the seed; among these, mol-miR166 (8 members), mol-miR167 (8 members), mol-miR156 (6 members) and mol-miR399 (6 members) were identified as well as 46 novel miRNAs.

Prediction of miRNA targets. Plant miRNAs play important roles in diverse biological processes by cleaving target mRNAs or suppressing their translation. In order to understand the biological functions of *M. oleifera* miRNAs, PNRD (plant non-coding RNA database)³⁶ was used as a reference for the prediction of putative target genes for tissue-specific and backbone miRNAs. The analysis revealed 7,677 unambiguous target genes for 54 miRNA families conserved across all the experimental conditions (Supplementary Table 5).

The majority of target genes were transcription factors; these include key regulators of abiotic and biotic stress response, plant reproduction and plant growth/development gene such as *squamosa promoter-binding proteins* (SPL or SBP)^{37,38}, *myb domain proteins* (MYBs)^{39–41}, *ARM repeat superfamily protein*^{42,43}, *Homeobox-Leucine Zipper Proteins* (ATHBs)^{44,45}, *Auxin response factors* (ARFs)^{46,47}, *Disease resistance protein*^{48,49}, *F-box family protein*⁵⁰ and *NAC domain transcription factors* (NACs)⁵¹.

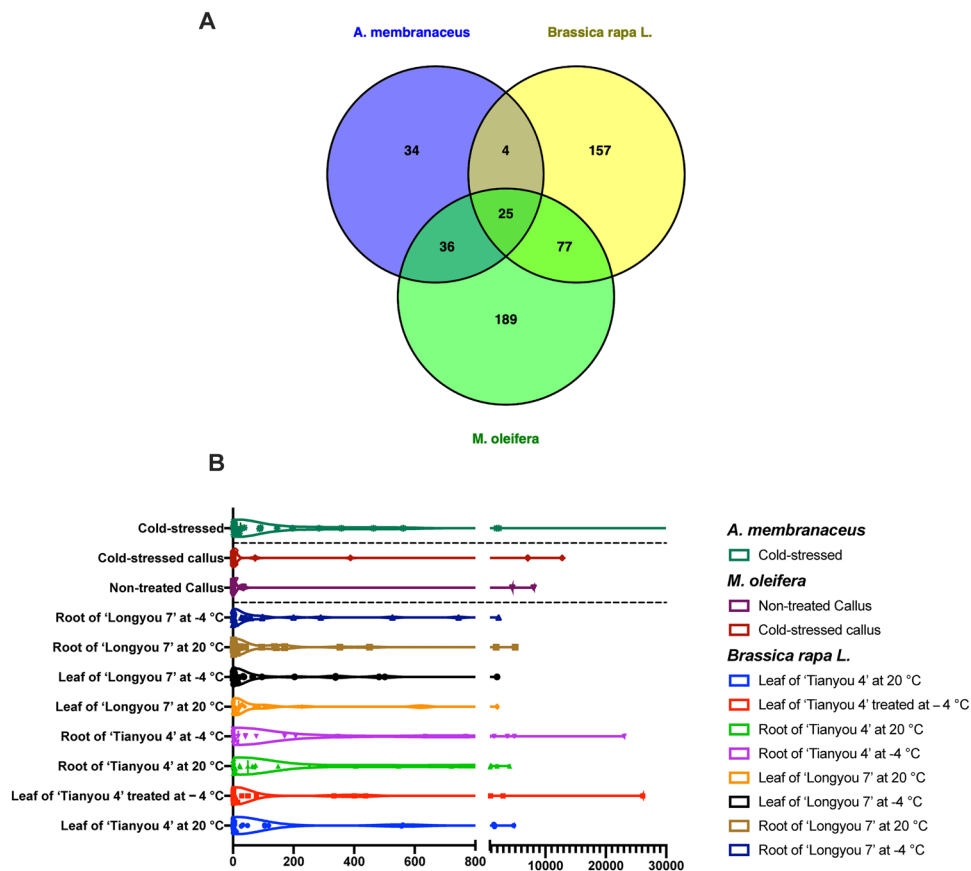


Figure 4. The number of miRNAs conserved across several cold-stress plants, is showed in a Venn diagram. (A) The mean abundance of this putative, “cold-responsive” miRNA signature across the different experimental conditions, is represented as a violin plot (B).

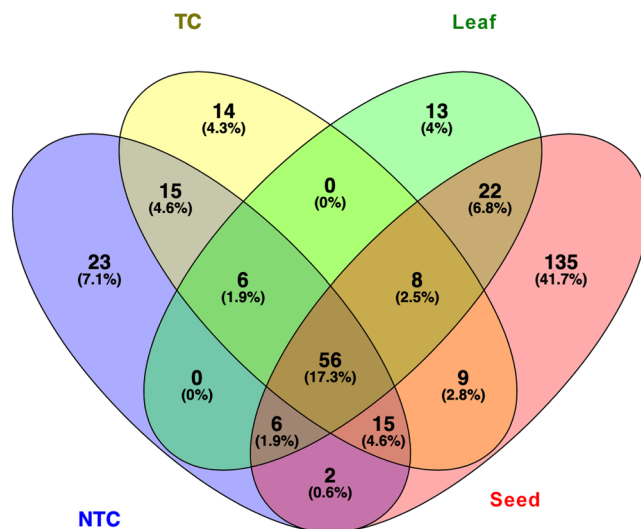


Figure 5. Venn diagram showing the number of conserved and novel miRNAs in accordance for all experimental conditions. NTC: Non Treated Callus. TC: Treated Callus.

Some of the identified miRNA targets are genes directly involved in resistance to biotic and abiotic stresses, including protein degradation (*ubiquitin carboxyl-terminal hydrolase 1*) and oxidoreductase activity (*ALDH22a1*).

Other predicted miRNA target genes are associated to plant growth/development, metabolism and plant reproduction; several genes are related to protein degradation (*auxin signaling f-box 2*), enzymatic activity

(laccase, RHOMBOID-like protein 1, clathrin heavy chain, zinc finger family protein, Hydrolase, Cysteine Protease, Cysteine/Histidine-rich C1 domain, Protein Kinase, NAD(P)-binding Rossmann-fold, Pectine lyase Kinase activity, RNA binding family protein, Methyltransferases and DEAD/DEAH box RNA helicase family protein), plant development (Pentatricopeptide repeat (PPR) superfamily protein, RING/U-box superfamily protein), translation process (Ribosomal protein) and oxidoreductase activity (2-oxoglutarate and Fe(II)-dependent oxygenase superfamily protein, Galactose oxidase).

Gene Ontology (GO) enrichment analysis. The enrichment analysis of miRNA target genes represents a fundamental procedure for evaluating the miRNA-mediated biological modulations that occur in *Moringa* leaves, seed, TC and NTC. The terms belonging to GO biological processes (BP), molecular functions (MF), and cellular compartments (CC) have been enriched and the results for the top 20 terms plotted in a bar plot (Supplementary Figures 1_NTC, TC, LEAF, SEED). Analyses performed on targets for backbone miRNAs, identified certain biological functions and molecular processes that are crucial for the proper survival of the plant. On the other hand, the enrichment analysis of targets for condition-specific miRNAs highlights biological modulations that are essential in a specific, biological context.

Biological processes. In the leaf, the majority of miRNAs seems to be responsible for fine regulation of genes involved in leaf symmetry^{52,53}; other genes regulate the development of different plant organs, such as anther⁵⁴, floral whorl⁵⁵, stamen⁵⁶ and so on. In callus, most miRNAs are involved in positive heterochronic gene regulation of development, especially in stress conditions; other genes are involved in the biosynthesis and metabolic processes of beta-D-glucan⁵⁷ (callus) and the lignin/phenylpropanoid catabolic processes⁵⁸. In the seed, most frequently expressed genes regulate the maintenance of DNA methylation, negative regulation of intracellular signal transduction and specification of axis polarity⁵⁹.

Cellular compartments. In leaf and seed, most of the regulations take place in the late endosome, vacuole and cell projection. In callus, most of the genes are localized in the 1.3-beta-D-glucan synthase complex (NTC) and mainly in the chloroplast endopeptidase Clp complex (TC). The presence of several genes located in the chloroplast could be correlated with the enrichment of stress-response genes. In fact, most of the transcriptional activities in the nucleus are regulated in part by signals derived from plastids. This process is named “retrograde signaling” and it is well amplified during responses to chemical, physical and biological stress⁶⁰.

Molecular functions. Despite most of the enriched terms do not provide a detailed description of the involved molecular function, they clearly highlight the fine-tuning role of microRNAs in plants. Indeed, most of the targets are involved in xenobiotic transmembrane transporting ATPase activity⁶¹ (leaf), oxidoreductase and ATP-dependent peptidase activity⁶² (TC), superoxide dismutase copper chaperone activity⁶³ (NTC), phosphorelay sensor kinase activity and xenobiotic transmembrane transporting ATPase activity⁶⁴ (seed).

Validation of conserved miRNAs by qRT-PCR. To confirm the expression pattern of the *M. oleifera* miRNAs, 11 conserved miRNAs with different expression profiles were randomly selected for real-time PCR analysis. The data collected demonstrate that the expression patterns are similar between the two analytical tools (Illumina sequencing and qRT-PCR) for six of the eleven miRNAs, whereas other show different expression pattern as detected between the two molecular tools.

As illustrated in Fig. 6, mol-miR396a and mol-miR159a are more abundant than other miRNAs in leaf and callus (both NTC and TC). Mol-miR162a is more abundantly expressed in leaf than in callus (both NTC and TC), while mol-miR398b and mol-miR168a are more expressed in calluses. On the other hand, mol-miR166i and mol-miR160h show similar levels of expression in all tissues.

Discussion

In the past decade, *M. oleifera* has gained growing attention for its nutraceutical and pharmacological functions as well as potential benefits for human health; several studies have demonstrated the great number of bioactive compounds contained in leaves, seeds, pods and flowers^{1,2,4}. The large number of these bioactive compounds (such as secondary metabolites) might explain the pharmacological properties of this “miracle tree”. Recent studies have confirmed the high content of polyphenols (such as flavonoids and phenolic acids) in *M. oleifera* leaves⁶⁵. These secondary metabolites are synthesized in the plant for its specialized need, in particular, ecological conditions; unlike primary metabolites, secondary metabolites are directly involved in the defence mechanism against environmental injuries. In human, *M. oleifera* polyphenols (such as flavonoids) are involved in protection against chronic diseases associated with oxidative stress, including cardiovascular disease and cancer⁴. Plant metabolic engineering has been used for enhancing biosynthesis of these pharmacological important phytochemicals⁶⁻⁹. One of the ways to increase polyphenol biosynthesis is by modulating levels of miRNAs – the ultimate regulators of biosynthesis and accumulation of secondary metabolites in plants⁶⁶⁻⁶⁸.

In our previous study²⁰ we reported a high throughput sequencing of miRNAs from *M. oleifera* seed; some of them were involved in regulation of human genes when transfected into Hepatoma cell lines (HepG2)⁶⁹. Being the part of the plant that contains the highest number of bioactive compounds, the leaf was chosen as model tissue for high-throughput sequencing in this study. Moreover, callus cells, given their undifferentiated nature, represent an optimal model for the identification and quantification of miRNA types present in *M. oleifera* in its basal condition. Low temperature was chosen as appropriate stress condition in the experimental procedure, considering that, due to natural geographical distribution of *M. oleifera*, this plant hardly ever comes into contact with this type of adverse condition. Indeed, cold stress has been shown to have an important influence on plant growth and development by modulation of gene expression. Despite a preliminary, explorative analysis did not highlight

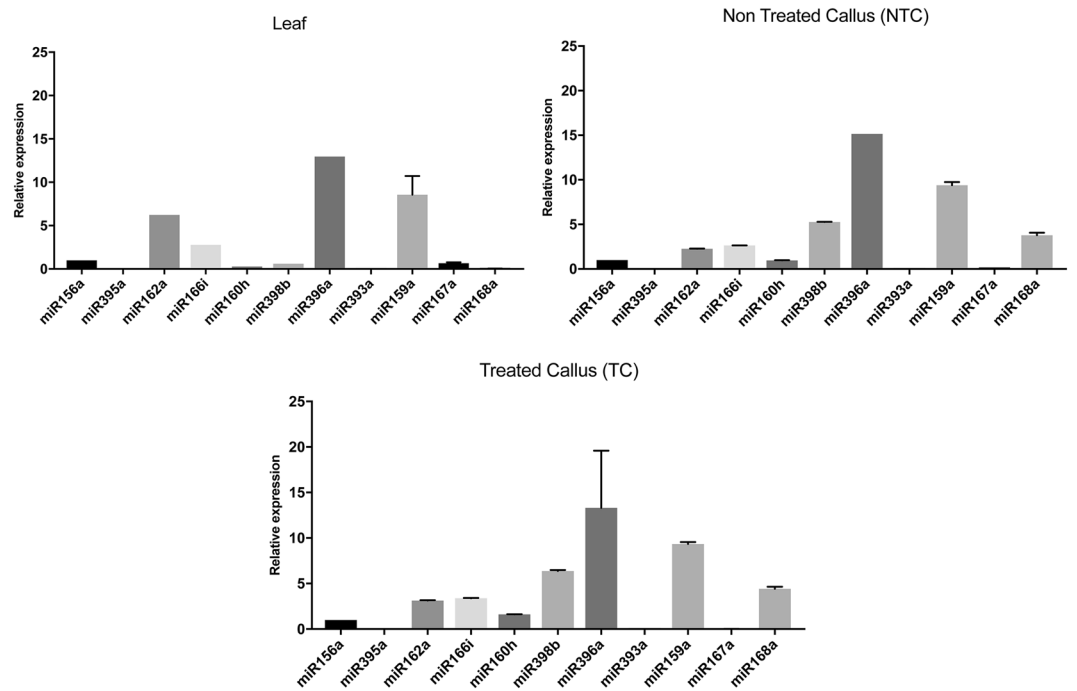


Figure 6. Quantitative RT-PCR analysis of conserved miRNAs in *M. oleifera* leaf and callus. The X axis represents different miRNAs. The Y axis represents the relative expression level of miRNAs. The 5S rRNA was used as an internal control. The expression level of miRNA156a was set as control and taken as 1 and expression level in all other miRNAs was quantified relative to it. The analysis was performed as triplicates, and the error bars indicate the standard deviations.

the presence of a “cold-sensitive” signature of microRNAs, specific isoforms were up-regulated in different plants. The application of the cold stress to a callus for further induction of somaclonal variations, could open, in turn, the possibility to develop a *M. oleifera* cell line better adapted to a temperate climate. For these reasons, *M. oleifera* calluses both untreated and exposed to abiotic stress seemed to be an interesting candidate models for this study.

The recent availability of data for the *M. oleifera* genome, allowed to re-evaluate our previous predictions from seeds and to identify a set of high-confidence and novel miRNAs from cold-stressed and non-stressed callus and leaf. A subsequent sequence alignment (Blast) compared to all plant-related mature miRNAs annotated in miRbase, allowed for identification of an additional set of low-confidence miRNAs. In the present study, a total of 131 known high-confidence miRNA isoforms (21 miRNA families), 300 known low-confidence miRNA isoforms (56 miRNA families) and 392 novel miRNA isoforms were identified in *M. oleifera* leaf, seed and callus with or without abiotic stress in the ten libraries. The majority of the identified miRNAs were 21 nt in length (60%; mean value of all experimental conditions), which is the canonical size for miRNAs generated from DCL1 processing. This result is similar to that of the conserved miRNAs predicted in *B. oleracea*⁷⁰ and other plant species such as soybean⁷¹, maize⁷², switchgrass⁷³ and Chinese cabbage^{74,75}. In our study, the number of reads of the conserved 21 nt miRNAs is very high, especially for the leaf (4.6×10^5 reads). The result was consistent with our previous reports in which sequence length 21nt has the highest number of miRNA reads in *M. oleifera* seeds.

Mol-miR166 and mol-miRNA156 families contain larger number of isomiRs in the seed, leaves and in the treated callus; these microRNAs are involved in many biological processes including leaf development⁷⁶, apical dominance⁷⁷, floral transition and development⁷⁸. MiRNA166, together with miRNA165, belong to another important class of miRNAs involved in *Shoot Apical Meristem (SAM)* maintenance. These two miRNAs share the same function of targeting/repressing *class III Homeodomain-Leucine Zipper (HD-ZIP III)* expression. These miRNAs play an important role in meristem maintenance, adaxial identity of leaves, lateral root growth and procambium identity.

The mol-miR156 family was a largely conserved miRNA family, followed by mol-miR395, mol-miR164 and mol-miR172. In our analysis, mol-miR156 family shows higher expression in leaf compared to other experimental conditions; these miRNAs belong to the largest families highly conserved in all land plants⁷⁹. Previous studies demonstrated their role in SPL pathway regulating plant development, flowering and plastochron length.

In plants, many miRNAs seem to be universally expressed among diverse species, such as miRNA156, miRNA157, miRNA159, miRNA160, miRNA161, miRNA171, and so on.

However, there is a large number of miRNAs present in only few other species. For instance, mol-miR6300 was discovered in soybean and subsequently identified in Finger millet and in Chickpea, as well as *M. oleifera* seed (as reported in our previous work²⁰). In flowers and leaves of Chickpea, this miRNA was found abundantly expressed; however, as predicted by other authors, its annotation in miRBase is ambiguous and needs experimental confirmation⁸⁰. In our analysis, all the experimental conditions (especially for leaf) showed an up-regulation

of this miRNA compared to other miRNAs. It would be interesting to investigate in more detail the biological function of this miRNA considering that it is highly expressed in many parts of this medicinal plant.

In this study, mol-miR159, mol-miR393 and mol-miR396 were found highly expressed in all experimental conditions. These results are in line with different studies, comprising various plant species, in which miRNA159 was shown to be very frequently detectable miRNA involved in fundamental plant biology roles, such as plant growth, and development⁸¹. In *A. thaliana*, miRNA159 has been shown to regulate anther and silique development by targeting MYB33. Mutation of the miRNA159-binding site on MYB33 mRNA resulted in pleiotropic defects including severely impaired fertility, stunted anthers, small siliques, and small seeds. Interestingly, miRNA159 accumulation was up-regulated by gibberellin (GA) application and GA-deficient mutants showed low miRNA159 accumulation. Treating these mutants with GA was itself sufficient to increase the accumulation of miRNA159 to wild-type levels and above, demonstrating the interplay between miRNAs and hormones in plant development⁸².

It is noteworthy to mention that miRNA159 is involved in cross-kingdom regulation of mammalian gene expression; Chin *et al.* found that plant miRNA159 could be detected in human sera and its levels were inversely correlated with breast cancer incidence and progression targeting human *Transcription Factor 7 (TCF7)* gene⁸³.

Gene targets were computationally predicted for both conserved and novel miRNAs in order to elucidate the biological functions of miRNAs in *M. oleifera* leaf, seed and callus. As demonstrated by other studies, mol-miR159, very common in all experimental conditions, was predicted to target *MYB* transcription factors; these target genes have been reported to play an important role in abscisic acid (ABA) signalling during *A. thaliana* seed germination⁸⁴. Moreover, mol-miR159 and its target gene (*GAMYB*) were involved in modulation of grapevine floral development in response to gibberellin (GA) treatments and this interaction has important implications for the molecular breeding of high-quality seedless grapevine berry. Mol-miR393, more common in seed than in leaf and callus, was shown to be involved in regulating various aspects of plant growth and development; for instance, mol-miR393 targets the *Transport Inhibitor Response 1 (TIR1)* and *Auxin Signaling F-Box (AFB)* genes, involved in flower development thus contributing to the maize grain filling rate by regulating maize growth, development and environment stress response⁸⁵. According to our analysis, this miRNA regulates enzymes involved mainly in Hydrolase and Methyltransferase activity including transcriptional regulation. *Growth-Regulating Factors (GRF)* genes are targeted and regulated by mol-miR396, very common in *M. oleifera* seed and leaf. In *A. thaliana* leaves in phase of development, expression of miRNA396 in the distal region of the leaf blade restricts GRF activity, thereby confining cell proliferation to the proximal leaf blade. As the leaf matures, miRNA396 level increases in the developed leaf leading to decreasing levels of GRF, stopping leaf blade growth. This conserved miRNA and its targets produced similar result in our analysis. Our results showed a very high conservation rate for mol-miR156 and a moderate abundance in *M. oleifera* leaf. This miRNA suppresses *SPL* and represses flowering in many plant species. As reported by various authors, the miRNA156 family is one of the largest miRNA gene families and members of the family are highly conserved in all land plants.

The predicted potential targets of *M. oleifera* novel miRNAs were involved in different molecular functions such as transcription factor activity, Auxin signalling and Ethylene responses. Many of these target genes are involved in plant growth and development, or in stress responses. Furthermore, our predicted analysis found a novel miRNA (mol-miR-n111) overexpressed in treated callus respect to non-treated callus that regulate many target genes, including *MYB* activity; previous reports have indicated that this transcription factors is also a target of miRNA159^{86,87}. In our analysis, mol-miR159 was up-regulated in all experimental conditions; in plants, MYBs are involved in different processes including primary and secondary metabolism, cell fate and identity, development, and responses to biotic and abiotic stresses including cold tolerance. The novel mol-miR-n323, commonly expressed in seed, was predicted to regulate different target genes, including *Histone Deacetylase 5 (HDA5)*. In *A. thaliana*, this protein is present in a protein complex involved in the regulation of flowering time⁸⁸; it's regulated by mol-miR482, uncommon in seed. In eukaryotes, epigenetic mechanisms (such as acetylation) play a crucial role in the regulation of gene expression; the HDACs protein complex has been associated with various developmental processes such as flowering, floral organ identity, seed development and circadian clocks.

In order to identify conserved and novel miRNAs involved in the biosynthesis of secondary metabolites in plants, we performed an overlaid analysis to identified specific miRNAs. Mol-miR408, more common in treated callus respect to others experimental conditions, are involved in the biosynthesis of benzyliso-quinoline alkaloids (BIA). This alkaloid is synthesized mainly by an agronomically and economically important medicinal plant, the Opium poppy (*Papaver somniferum* L.), used for main morphine alkaloid production⁶⁷. In *Argemone mexicana* L. (Papaveraceae), BIA is used to treat different disorders, given its antimicrobial, antiparasitic, antimalarial, pesticide, cytotoxic and neurological properties⁸⁹. Another important miRNA implicated in the biosynthesis of the secondary metabolites in plants is mol-miR156. This miRNA, targeting *SPL9* gene, is up-regulated in *M. oleifera* leaf and seed and seems to be involved in the Anthocyanin and Sesquiterpenoid/triterpenoid biosynthesis. Anthocyanins are flavonoids with antioxidant properties and can therefore can be used potentially as dietary nutraceuticals for human health. Data from several epidemiological studies have reported an inverse correlation between anthocyanin intake and risk of cardiovascular disease (CVD) or CVD-related mortality⁹⁰. Different authors have shown a beneficial effect of anthocyanin rich foods on gut microbiota; the result of *in vitro* and *in vivo* studies highlighted a significant proliferative effect on *Bifidobacterium* spp., known for their wide use in probiotics and for the treatment of irritable bowel syndrome, as well as inhibition of *Clostridium histolyticum*, pathogenic in humans⁹¹. Mol-miR396b_1, upregulated in *M. oleifera* seed and leaf, is well known for being able to regulate the biosynthesis of Flavonol glycoside by target Kaempferol 3-O-beta-D-galactosyltransferase⁹². Dietary flavonoids isolated from different medicinal plants have received an increased attention due to their considerable benefits in the prevention and management of modern diseases such as cancers, diabetes, and cardiovascular diseases⁹³.

In conclusion, this is the first comprehensive identification of conserved and novel miRNAs in *M. oleifera* leaf, seed and callus (stressed and non-stressed). This dataset represents an important supplement to the existing *M. oleifera* miRNA database and shall be useful in understanding the possible role for enhancing biosynthesis of pharmacologically important phytochemicals by plant miRNAs during biotic and abiotic stresses. Further studies are necessary in order to elucidate this complex regulatory network potentially able to improve human health in socially neglected populations by a “miracle” tree with high nutritional and medicinal value.

Data availability

The sequences from the small RNA library have been deposited in the Gene Expression Omnibus (GEO) database, the accession number is GSE119247 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE119247>).

Received: 19 March 2019; Accepted: 20 September 2019;

Published online: 22 October 2019

References

- Anwar, F., Latif, S., Ashraf, M. & Gilani, A. H. Moringa oleifera: a food plant with multiple medicinal uses. *Phytotherapy Research* **21**, 17–25 (2006).
- Sánchez-Machado, D. I., Núñez-Gastélum, J. A., Reyes-Moreno, C., Ramírez-Wong, B. & López-Cervantes, J. Nutritional Quality of Edible Parts of Moringa oleifera. *Food Analytical Methods* **3**, 175–180 (2009).
- Jung, I. L. Soluble Extract from Moringa oleifera Leaves with a New Anticancer Activity. *PLoS ONE* **9**, e95492 (2014).
- Vergara-Jimenez, M., Almatrafi, M. & Fernandez, M. Bioactive Components in Moringa Oleifera Leaves Protect against Chronic Disease. *Antioxidants* **6**, 91 (2017).
- Vlahov, G., Chepkwony, P. K. & Ndalut, P. K. ¹³C NMR Characterization of Triacylglycerols of Moringa oleifera Seed Oil: An “Oleic-Vaccenic Acid” Oil. *Journal of Agricultural and Food Chemistry* **50**, 970–975 (2002).
- Moscattello, R., Baldan, B. & Navazio, L. Plant Cell Suspension Cultures. *Plant Mineral Nutrients* **77–93**, https://doi.org/10.1007/978-1-62703-152-3_5 (2012).
- Tabata, H. Paclitaxel Production by Plant-Cell-Culture Technology. *Biomufacturing* **1–23**, <https://doi.org/10.1007/b13538> (2004).
- Grech-Baran, M., Sykłowska-Baranek, K. & Pietrosiuk, A. Biotechnological approaches to enhance salidroside, rosin and its derivatives production in selected *Rhodiola* spp. *in vitro* cultures. *Phytochemistry Reviews* **14**, 657–674 (2014).
- Wróbel, T., Dreger, M., Wielgus, K. & Słomski, R. The application of plant *in vitro* cultures in cannabinoid production. *Biotechnology Letters* **40**, 445–454 (2017).
- Bartel, D. P. MicroRNAs: Target Recognition and Regulatory Functions. *Cell* **136**, 215–233 (2009).
- Axtell, M. J., Westholm, J. O. & Lai, E. C. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. *Genome Biology* **12**, 221 (2011).
- Rogers, K. & Chen, X. microRNA Biogenesis and Turnover in Plants. *Cold Spring Harbor Symposia on Quantitative Biology* **77**, 183–194 (2012).
- Budak, H. & Akpınar, B. A. Plant miRNAs: biogenesis, organization and origins. *Functional & Integrative Genomics* **15**, 523–531 (2015).
- Gupta, O. P., Meena, N. L., Sharma, I. & Sharma, P. Differential regulation of microRNAs in response to osmotic, salt and cold stresses in wheat. *Molecular Biology Reports* **41**, 4623–4629 (2014).
- Shriram, V., Kumar, V., Devarumath, R. M., Khare, T. S. & Wani, S. H. MicroRNAs As Potential Targets for Abiotic Stress Tolerance in Plants. *Frontiers in Plant Science* **7** (2016).
- Zhang, Y. *et al.* Identification and characterization of cold-responsive microRNAs in tea plant (*Camellia sinensis*) and their targets using high-throughput sequencing and degradome analysis. *BMC Plant Biol.* **14**, 271 (2014).
- Cao, X., Wu, Z., Jiang, F., Zhou, R. & Yang, Z. Identification of chilling stress-responsive tomato microRNAs and their target genes by high-throughput sequencing and degradome analysis. *BMC Genomics* **15**, 1130 (2014).
- Yang, X., Liu, F., Zhang, Y., Wang, L. & Cheng, Y.-F. Cold-responsive miRNAs and their target genes in the wild eggplant species *Solanum aculeatissimum*. *BMC Genomics* **18**, 1000 (2017).
- Zhou, X., Liu, J., Bao, S., Yang, Y. & Zhuang, Y. Molecular Cloning and Characterization of a Wild Eggplant *Solanum aculeatissimum* NBS-LRR Gene, Involved in Plant Resistance to *Meloidogyne incognita*. *Int J Mol Sci* **19**, (2018).
- Pirró, S. *et al.* MicroRNA from Moringa oleifera: Identification by High Throughput Sequencing and Their Potential Contribution to Plant Medicinal Value. *PLoS ONE* **11**, e0149495 (2016).
- Tian, Y. *et al.* High quality reference genome of drumstick tree (*Moringa oleifera* Lam.), a potential perennial crop. *Sci China Life Sci* **58**, 627–638 (2015).
- Murashige, T. & Skoog, F. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* **15**, 473–497 (1962).
- Andrews, S. FastQC A Quality Control tool for High Throughput Sequence Data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* **17**, 10 (2011).
- Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res.* **40**, 37–52 (2012).
- Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
- Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research* **42**, D68–D73 (2013).
- Dai, X., Zhuang, Z. & Zhao, P. X. psRNA Target: a plant small RNA target analysis server (2017 release). *Nucleic Acids Res.* **46**, W49–W54 (2018).
- Mi, H. *et al.* PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* **45**, D183–D189 (2017).
- Gismondì, A., Di Marco, G. & Canini, A. Detection of plant microRNAs in honey. *PLOS ONE* **12**, e0172981 (2017).
- Zhou, P. *et al.* Integrated analysis of transcriptomic and metabolomic data reveals critical metabolic pathways involved in polyphenol biosynthesis in *Nicotiana tabacum* under chilling stress. *Funct. Plant Biol.* **46**, 30–43 (2018).
- Zhang, C. *et al.* Functional conservation analysis and expression modes of grape anthocyanin synthesis genes responsive to low temperature stress. *Gene* **574**, 168–177 (2015).
- Zeng, X. *et al.* Identification of cold stress responsive microRNAs in two winter turnip rape (*Brassica rapa* L.) by high throughput sequencing. *BMC Plant Biology* **18** (2018).
- Abla, M. *et al.* Identification of miRNAs and Their Response to Cold Stress in *Astragalus Membranaceus*. *Biomolecules* **9** (2019).
- Ivashuta, S. *et al.* Regulation of gene expression in plants through miRNA inactivation. *PLoS ONE* **6**, e21330 (2011).
- Yi, X., Zhang, Z., Ling, Y., Xu, W. & Su, Z. PNRD: a plant non-coding RNA database. *Nucleic Acids Research* **43**, D982–D989 (2014).

37. Chen, X. *et al.* SQUAMOSA promoter-binding protein-like transcription factors: star players for plant growth and development. *J Integr Plant Biol* **52**, 946–951 (2010).
38. Zhou, Q. *et al.* Genome-wide identification and characterization of the SBP-box gene family in Petunia. *BMC Genomics* **19**, 193 (2018).
39. Millar, A. A. & Gubler, F. The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell* **17**, 705–721 (2005).
40. Gocal, G. F. *et al.* GAMYB-like genes, flowering, and gibberellin signaling in Arabidopsis. *Plant Physiol.* **127**, 1682–1693 (2001).
41. Zhu, N. *et al.* The R2R3-type MYB gene OsMYB91 has a function in coordinating plant growth and salt stress tolerance in rice. *Plant Sci.* **236**, 146–156 (2015).
42. Sharma, M. & Pandey, G. K. Expansion and Function of Repeat Domain Proteins During Stress and Development in Plants. *Front Plant Sci* **6** (2016).
43. Sharma, M. *et al.* Comprehensive Expression Analysis of Rice Armadillo Gene Family During Abiotic Stress and Development. *DNA Res* **21**, 267–283 (2014).
44. Aoyama, T. *et al.* Ectopic expression of the Arabidopsis transcriptional activator Athb-1 alters leaf cell fate in tobacco. *Plant Cell* **7**, 1773–1785 (1995).
45. Henriksson, E. *et al.* Homeodomain leucine zipper class I genes in Arabidopsis. Expression patterns and phylogenetic relationships. *Plant Physiol.* **139**, 509–518 (2005).
46. Bouzroud, S. *et al.* Auxin Response Factors (ARFs) are potential mediators of auxin action in tomato response to biotic and abiotic stress (*Solanum lycopersicum*). *PLoS ONE* **13**, e0193517 (2018).
47. Yu, C., Zhan, Y., Feng, X., Huang, Z.-A. & Sun, C. Identification and Expression Profiling of the Auxin Response Factors in *Capsicum annuum* L. under Abiotic Stress and Hormone Treatments. *Int J Mol Sci* **18** (2017).
48. Martin, G. B., Bogdanove, A. J. & Sessa, G. Understanding the functions of plant disease resistance proteins. *Annu Rev Plant Biol* **54**, 23–61 (2003).
49. Belkadir, Y., Subramaniam, R. & Dangl, J. L. Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Curr. Opin. Plant Biol.* **7**, 391–399 (2004).
50. Xu, G., Ma, H., Nei, M. & Kong, H. Evolution of F-box genes in plants: Different modes of sequence divergence and their relationships with functional diversification. *PNAS* **106**, 835–840 (2009).
51. Nuruzzaman, M., Sharoni, A. M. & Kikuchi, S. Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants. *Front Microbiol* **4** (2013).
52. Moubayidin, L. & Østergaard, L. Dynamic Control of Auxin Distribution Imposes a Bilateral-to-Radial Symmetry Switch during Gynoecium Development. *Curr Biol* **24**, 2743–2748 (2014).
53. Gao, R., Wang, Y., Gruber, M. Y. & Hannoufa, A. miR156/SPL10 Modulates Lateral Root Development, Branching and Leaf Morphology in Arabidopsis by Silencing AGAMOUS-LIKE 79. *Front Plant Sci* **8**, 2226 (2017).
54. Ogawa, M., Kay, P., Wilson, S. & Swain, S. M. Arabidopsis dehiscence zone polygalacturonase1 (ADPG1), ADPG2, and QUARTET2 are Polygalacturonases required for cell separation during reproductive development in Arabidopsis. *Plant Cell* **21**, 216–233 (2009).
55. Hong, Y. & Jackson, S. Floral induction and flower formation—the role and potential applications of miRNAs. *Plant Biotechnol. J.* **13**, 282–292 (2015).
56. Song, S., Qi, T., Huang, H. & Xie, D. Regulation of Stamen Development by Coordinated Actions of Jasmonate, Auxin, and Gibberellin in Arabidopsis. *Molecular Plant* **6**, 1065–1073 (2013).
57. Burton, R. A. *et al.* Cellulose synthase-like CslF genes mediate the synthesis of cell wall (1,3;1,4)-beta-D-glucans. *Science* **311**, 1940–1942 (2006).
58. Yoon, J., Choi, H. & An, G. Roles of lignin biosynthesis and regulatory genes in plant development. *Journal of Integrative Plant Biology.* <https://doi.org/10.1111/jipb.12422> (2015).
59. Bartels, A. *et al.* Dynamic DNA Methylation in Plant Growth and Development. *Int J Mol Sci* **19** (2018).
60. Crawford, T., Lehotai, N. & Strand, Å. The role of retrograde signals during plant stress responses. *J. Exp. Bot.* **69**, 2783–2795 (2018).
61. Bernard, C. *et al.* Equilibrative nucleoside transporter 1 (ENT1) is critical for pollen germination and vegetative growth in Arabidopsis. *J Exp Bot* **62**, 4627–4637 (2011).
62. Ikeuchi, M. *et al.* Wounding Triggers Callus Formation via Dynamic Hormonal and Transcriptional Changes. *Plant Physiology* **175**, 1158–1174 (2017).
63. Ślesak, I., Libik, M. & Miszalski, Z. Superoxide dismutase activity in callus from the C3-CAM intermediate plant *Mesembryanthemum crystallinum*. *Plant Cell, Tissue and Organ Culture* **75**, 49–55 (2003).
64. Bártíková, H. *et al.* Xenobiotic-metabolizing enzymes in plants and their role in uptake and biotransformation of veterinary drugs in the environment. *Drug Metab. Rev.* **47**, 374–387 (2015).
65. Verma, A. R., Vijayakumar, M., Mathela, C. S. & Rao, C. V. *In vitro* and *in vivo* antioxidant properties of different fractions of *Moringa oleifera* leaves. *Food and Chemical Toxicology* **47**, 2196–2201 (2009).
66. Singh, N. & Sharma, A. Turmeric (*Curcuma longa*): miRNAs and their regulating targets are involved in development and secondary metabolite pathways. *Comptes Rendus Biologies* **340**, 481–491 (2017).
67. Boke, H. *et al.* Regulation of the alkaloid biosynthesis by miRNA in opium poppy. *Plant Biotechnology Journal* **13**, 409–420 (2015).
68. Robert-Seilaniantz, A. *et al.* The microRNA miR393 re-directs secondary metabolite biosynthesis away from camalexin and towards glucosinolates. *The Plant Journal* **67**, 218–231 (2011).
69. Pirrò, S. *et al.* Bioinformatics Prediction and Experimental Validation of MicroRNAs Involved in Cross-Kingdom Interaction. *J. Comput. Biol.* **23**, 976–989 (2016).
70. Lukasił, A., Pietrykowska, H., Paczek, L., Szweykowska-Kulinska, Z. & Zielenkiewicz, P. High-throughput sequencing identification of novel and conserved miRNAs in the Brassica oleracea leaves. *BMC Genomics* **14**, 801 (2013).
71. Zhang, Q. *et al.* Association of the circadian rhythmic expression of GmCRY1a with a latitudinal cline in photoperiodic flowering of soybean. *Proceedings of the National Academy of Sciences* **105**, 21028–21033 (2008).
72. Zhang, B., Pan, X. & Anderson, T. A. Identification of 188 conserved maize microRNAs and their targets. *FEBS Letters* **580**, 3753–3762 (2006).
73. Xie, F., Frazier, T. P. & Zhang, B. Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). *Planta* **232**, 417–434 (2010).
74. Dhandapani, V. *et al.* Identification of potential microRNAs and their targets in Brassica rapa L. *Molecules and Cells* **32**, 21–37 (2011).
75. Wang, J., Hou, X. & Yang, X. Identification of conserved microRNAs and their targets in Chinese cabbage (*Brassica rapa* subsp. *pekinensis*). *Genome* **54**, 1029–1040 (2011).
76. Rubio-Somoza, I. *et al.* Temporal Control of Leaf Complexity by miRNA-Regulated Licensing of Protein Complexes. *Current Biology* **24**, 2714–2719 (2014).
77. Chen, Z., Gao, X. & Zhang, J. Alteration of osa-miR156e expression affects rice plant architecture and strigolactones (SLs) pathway. *Plant Cell Reports* **34**, 767–781 (2015).
78. Li, C. & Zhang, B. MicroRNAs in Control of Plant Development. *Journal of Cellular Physiology* **231**, 303–313 (2015).
79. D’Ario, M., Griffiths-Jones, S. & Kim, M. Small RNAs: Big Impact on Plant Development. *Trends Plant Sci.* **22**, 1056–1068 (2017).

80. Srivastava, S. *et al.* High throughput sequencing of small RNA component of leaves and inflorescence revealed conserved and novel miRNAs as well as phasiRNA loci in chickpea. *Plant Science* **235**, 46–57 (2015).
81. Djami-Tchatchou, A. T., Sanan-Mishra, N., Ntushelo, K. & Dubery, I. A. Functional Roles of microRNAs in Agronomically Important Plants—Potential as Targets for Crop Improvement and Protection. *Frontiers in Plant Science* **8** (2017).
82. Achard, P., Herr, A., Baulcombe, D. C. & Harberd, N. P. Modulation of floral development by a gibberellin-regulated microRNA. *Development* **131**, 3357–3365 (2004).
83. Chin, A. R. *et al.* Cross-kingdom inhibition of breast cancer growth by plant miR159. *Cell Research* **26**, 217–228 (2016).
84. Chen, H. *et al.* Integration of light and abscisic acid signaling during seed germination and early seedling development. *Proceedings of the National Academy of Sciences* **105**, 4495–4500 (2008).
85. Jin, X. *et al.* Identification and Characterization of microRNAs during Maize Grain Filling. *PLoS ONE* **10**, e0125800 (2015).
86. Huang, J.-H. *et al.* Illumina microRNA profiles reveal the involvement of miR397a in Citrus adaptation to long-term boron toxicity via modulating secondary cell-wall biosynthesis. *Scientific Reports* **6** (2016).
87. Jiu, S. *et al.* Identification of copper (Cu) stress-responsive grapevine microRNAs and their target genes by high-throughput sequencing. *R Soc Open Sci* **6**, 180735 (2019).
88. Luo, M. *et al.* Regulation of flowering time by the histone deacetylase HDA5 in Arabidopsis. *Plant J.* **82**, 925–936 (2015).
89. Rubio-Pina, J. & Vazquez-Flota, F. Pharmaceutical applications of the benzyloisoquinoline alkaloids from *Argemone mexicana* L. *Curr Top Med Chem* **13**, 2200–2207 (2013).
90. Krga, I. & Milenkovic, D. Anthocyanins: From Sources and Bioavailability to Cardiovascular-Health Benefits and Molecular Mechanisms of Action. *J. Agric. Food Chem.* **67**, 1771–1783 (2019).
91. Igwe, E. O., Charlton, K. E., Probst, Y. C., Kent, K. & Netzel, M. E. A systematic literature review of the effect of anthocyanins on gut microbiota populations. *J Hum Nutr Diet* **32**, 53–62 (2019).
92. Prakash, P., Rajakani, R. & Gupta, V. Transcriptome-wide identification of *Rauvolfia serpentina* microRNAs and prediction of their potential targets. *Comput Biol Chem* **61**, 62–74 (2016).
93. Xiao, J., Capanoglu, E., Jassbi, A. R. & Miron, A. Advance on the Flavonoid C-glycosides and Health Benefits. *Crit Rev Food Sci Nutr* **56**(Suppl 1), S29–45 (2016).

Acknowledgements

We thank SOCOPOMO S/C 46P.A. de Lingang, Dschang, Cameroon for providing plant material. We thank Genomix4Life for sequence experiments.

Author contributions

S.P. designed the study, carried out the bioinformatics analysis, interpreted the results and wrote the manuscript. I.M. designed the study, wrote the manuscript and contributed to manuscript revising. A. Gu. and L.Z. performed the experiments and wrote the manuscript. A. Gi. performed the experiments and contributed to manuscript revising. R.C. and R.B. contributed reagents/materials/revised manuscript. V.C., A.C. and M.M. contributed and supervised the workflow of the project. A. Ga. designed the study, interpreted the results and wrote the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-51100-4>.

Correspondence and requests for materials should be addressed to G.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019