Bioinformatic Identification and Expression Analysis of Banana MicroRNAs and Their Targets

Juan Chai1☯, Renjun Feng1☯, Hourui Shi1,2, Mengyun Ren1,2, Yindong Zhang2*, Jingyi Wang1*

1 Key Laboratory of Biology and Genetic Resources of Tropical Crops, Ministry of Agriculture, Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou, Hainan, China, 2 College of Agronomy, Hainan University, Haikou, Hainan, China

☯ These authors contributed equally to this work.
* wangjingyi@itbb.org.cn (JYW), 23900598@163.com (YDZ)

Abstract

MicroRNAs (miRNAs) represent a class of endogenous non-coding small RNAs that play important roles in multiple biological processes by degrading targeted mRNAs or repressing mRNA translation. Thousands of miRNAs have been identified in many plant species, whereas only a limited number of miRNAs have been predicted in *M. acuminata* (A genome) and *M. balbisiana* (B genome). Here, previously known plant miRNAs were BLAST-ed against the Expressed Sequence Tag (EST) and Genomic Survey Sequence (GSS), a database of banana genes. A total of 32 potential miRNAs belonging to 13 miRNAs families were detected using a range of filtering criteria. 244 miRNA:target pairs were subsequently predicted, most of which encode transcription factors or enzymes that participate in the regulation of development, growth, metabolism, and other physiological processes. In order to validate the predicted miRNAs and the mutual relationship between miRNAs and their target genes, qRT-PCR was applied to detect the tissue-specific expression levels of 12 putative miRNAs and 6 target genes in roots, leaves, flowers, and fruits. This study provides some important information about banana pre-miRNAs, mature miRNAs, and miRNA target genes and these findings can be applied to future research of miRNA functions.

Introduction

MicroRNAs (miRNAs) are a class of endogenous non-coding RNA molecules that are 21–24 nucleotides (nt) in length. The processing and maturation of miRNAs in plants involves several steps that employ key enzymes, such as Dicer-like 1 (DCL1) [1] and HASTY [2,3]. Mature miRNAs are incorporated into the RNA-induced silencing complex (RISC) to post-transcriptionally regulate gene expression by targeting mRNA degradation or translation repression [4–6]. Previous research has shown that miRNAs play a key role in various metabolic and biological processes in plants, including meristem cell identity, leaf morphogenesis and polarity, floral
differentiation and development, hormone signaling, organ separation, fruit growth and development, and the stress responses [7].

A number of methods have been applied to miRNA prediction, including genetic screening, direct cloning, sequencing methods, and computational approaches. Compared to these methods, the computational approach is one of the most effective, rapid, and affordable strategies to identify miRNAs [8] based on the major characteristic features of miRNAs, such as the hairpin-shaped stem-loop secondary structure with minimal folding free energy and high evolutionary conservation between species [9–11]. Using bioinformatics analysis based on comparative genome-based homolog searching, expressed sequence tag (EST) and/or genomic survey sequence (GSS), many researchers have identified miRNA genes from a variety of species, including *Arabidopsis thaliana* [12], *Oryza sativa* [13], *Triticum aestivum* [14], *Zea mays* [15], *Hordeum vulgare* [16], *Sorghum bicolor* [17], *Gossypium* [18], *Glycine max* [19], *solanum tuberosum* [20], *Lycopersicon esculentum* [21], *Fragaria vesca* [7], *Malus domestica* [23], *Medicago truncatula* [24], *Brassica napus* [25], and *Catharanthus roseus* [26].

Bananas (*Musa* spp.) are giant perennial herbs from the order Zingiberales of the monocotyledon group, which originated in Southeast Asia and the western Pacific [27,28]. Cultivated bananas have evolved from the hybridization of wild species of *M. acuminata* and *M. balbisiana* [29]. These crops are of extreme importance across the world’s tropical and subtropical regions where they are vital for food security and are also a major source of incomes. Bananas are widely exported to industrialized countries where they constitute a major staple food for millions of people [27]. Bananas are the world’s most important fruit and rank fourth in the list of global food crops, with a global annual production in excess of 107 million tons [30].

Thousands of miRNAs belonging to various plant species have been registered in the miRNA database, miRBase [31] (http://www.mirbase.org), and novel miRNA genes will be identified in Release 20; however, none of these miRNAs are from *Musa*. Recently, a reference sequence of *Musa* A-genome and a draft sequence of B-genome have been published [27,32]. D’Hont et al (2012) reported 235 conserved miRNAs from 37 families using computational methods based on the Pahang A-genome sequence similarity with known miRNAs in PMRD (http://bioinformatics.cau.edu.cn/PMRD/). Only the miR528 family out of 8 poales genomes-specific miRNA families was found in the A-genome. Davey et al (2013) predicted a similar number of microRNA sequences (47 and 42 miRNA families) in both A- and B-genomes, respectively, using publically available small RNA data from *M. acuminata* Calcutta 4 leaf, flower, and fruit tissues (sequenced within the framework of a NSF project http://smallrna.udel.edu) and 235 miRNA sequences reported for A-genome [27]. Davey et al (2013) also identified 28 novel miRNAs that were unique to either *Musa* A- or B-genomes, whereas, they did not experimentally verify and characterize the miRNAs.

In the present study, a computational analysis and a homolog search were used to identify banana miRNAs using miRBase, as well as EST and GSS data in National Center for Biotechnology Information (NCBI). The potential target genes of miRNAs were identified, and the potential functions of the predicted miRNAs targets, mainly the transcription factor genes, were investigated. Using real time fluorescent quantitative reverse transcription polymerase chain reaction (qRT-PCR), selected miRNAs and their targets were used to analyze the expression patterns in different tissues (i.e., roots, leaves, flowers, and fruits). The identification of miRNAs and their targets is important in order to understand the roles of miRNAs in the growth development and stress physiology of the banana (*Musa* spp.).
Materials and Methods

Plant material

Roots, leaves, flowers, and fruits were collected from banana plants (Musa acuminata L. AAA group, ‘Brazilian’) grown at the fruit science experimental station of Chinese Academy of Tropical Agricultural Sciences (19°98′N, 110°33′E; Haikou, Hainan province, China). After collection, all the samples were immediately frozen in liquid nitrogen and stored at -80°C until use.

Sequences

A total of previously known plant miRNA sequences were obtained from miRBase [31] (http://www.mirbase.org/, Release 20.0, June 2013). In order to avoid overlap of sequences, repeats within the above sequences were removed. The remaining miRNA sequences were defined as a reference to find the conserved miRNAs in banana. A total of 46,111 banana EST and 31,544 GSS were obtained from the NCBI GenBank [33] in October 2013 (http://www.ncbi.nlm.nih.gov). These sequences were then used to predict the miRNAs and their target genes candidates.

Prediction of miRNAs and their precursors

The prediction procedure is shown in Fig 1. First, known plant miRNAs in the miRBase were used as query sequences to blast against the downloaded ESTs and GSSs. The EST and GSS sequences that had no more than 3 mismatching nucleotides against the known miRNAs were regarded as the homologous sequences of the corresponding miRNAs, and the franking regions plus the matching nucleotide sequences were taken as the potential miRNA precursors. The repeat sequences and those coding proteins were then removed from the hit sequence using BLAST 2.2.14 (ftp://ftp.ncbi.nih.gov/blast/) and BLASTX (http://www.ncbi.nlm.nih.gov/BLAST/). Analysis of secondary structure and free energy of these sequences was conducted using the web-based software Mfold 3.2 [34] (http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form).

These criteria were used for screening potential miRNAs or pre-miRNAs [35,36]: (1) predicted mature miRNAs have no more than 3 nt mismatches with the known miRNAs, (2) pre-miRNA sequence can fold into an appropriate hairpin secondary structure and the mature miRNA is located in one arm of the hairpin structure, (3) the mature miRNA sequence has less than 6 mismatches with the opposite miRNA/C3 sequence in the other arm and no loop or break in miRNA/miRNA/C3 duplex, (4) A+U content of pre-miRNA is ranged from 30% to 70%, and (5) the predicted secondary structure must have lower minimal folding free energy (MFE≤-20 kcal/mol).

Prediction of the potential miRNA targets

Based on the newly identified potential miRNA sequences, we also predicted their potential targets using online software psRNATarget [37] (http://plantgrn.noble.org/psRNATarget/) with the following default parameters: maximum expectation of 3.0, length for complementarity scoring (hspsize) of 20 bp, target accessibility-allowed maximum energy to unpair the target site (UPE) of 25.0, flanking length around target site for target accessibility analysis of 17 bp in upstream and 13 bp in downstream, and range of central mismatch leading to translation inhibition of 9–11 nt.

Expression analysis of banana miRNAs and target genes

Total RNA was isolated from different tissues following the CTAB method [38] and briefly exposed to RNAase-free DNAase I. cDNAs were synthesized by M-MLV according to the
manufacturer’s instruction, using the specific stem-loop primers for miRNAs or Oligo(dT)15 for their target mRNAs. The specific stem-loop primers for miRNAs are listed in Table 1. After reverse transcription, the products of each reaction were diluted 10 times to avoid potential primer interference in the qRT-PCR reactions.

The qRT-PCR reactions were performed using FastStart Universal SYBR Green Master (ROX) on a Mx3005P Real-Time PCR System. Briefly, 2.0 μL of diluted cDNA template was added to 12.5 μL of the FastStart Universal SYBR Green Master (ROX), 0.8 μM of each primer and ddH2O to a final volume of 25 μL. The reactions were amplified for 3 m at 95°C, followed by 35 or 40 cycles of 95°C for 10 s and 50°C for 30 s. A melting curve analysis was performed to determine the specificity of the products. Reactions were done in triplicate and the controls with no template and no reverse transcription were included for each gene. The banana 5S rRNA was used as an internal control. The forward and reverse primers used in real-time PCR are shown in Table 1 and Table 2.

The relative expression quantification was calculated using the equation $2^{-\Delta\Delta Ct}$, where $\Delta\Delta Ct = (Ct_{miRNA} - Ct_{5S rRNA})_{Tissue X} - (Ct_{miRNA} - Ct_{5S rRNA})_{Tissue 1}$ [39]. In this study, tissue 1 presents the expression level of each miRNA in leaves, and the expression levels
### Table 1. Stem-loop RT primers and PCR primers for 12 identified miRNAs.

| Name     | Sequence(5'-3')                                                                 | PCR prime Forward/Reverse                      |
|----------|--------------------------------------------------------------------------------|-------------------------------------------------|
| miR156d  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggtgctctc              |                                                 |
|          | PCR prime: Forward: acactccagctgggttgacagaag                                     | Reverse: aactggtgtcgtggag                       |
| miR156g  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggtgctctc              |                                                 |
|          | PCR prime: Forward: acactccagctgggttgacagaag                                     | Reverse: aactggtgtcgtggag                       |
| miR166b  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgagccccgagc              |                                                 |
|          | PCR prime: Forward: acactccagctgggqaaggtgt                                     | Reverse: aactggtgtcgtggag                       |
| miR399a  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggcagggcaa             |                                                 |
|          | PCR prime: Forward: acactccagctgggtgccaaagga                                     | Reverse: aactggtgtcgtggag                       |
| miR319m  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgagagggagct              |                                                 |
|          | PCR prime: Forward: acactccagctgggtggagaggtgt                                   | Reverse: aactggtgtcgtggag                       |
| miR160a  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggtgctgtg              |                                                 |
|          | PCR prime: Forward: acactccagctgggtgccaaagga                                     | Reverse: aactggtgtcgtggag                       |
| miR164e  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggtgctgtg              |                                                 |
|          | PCR prime: Forward: acactccagctgggtgccaaagga                                     | Reverse: aactggtgtcgtggag                       |
| miR167c  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgagtcagatcat             |                                                 |
|          | PCR prime: Forward: acactccagctgggtggagaggtgt                                   | Reverse: aactggtgtcgtggag                       |
| miR169h  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggtgctgtg              |                                                 |
|          | PCR prime: Forward: acactccagctgggtgccaaagga                                     | Reverse: aactggtgtcgtggag                       |
| miR162   | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggtgctgtg              |                                                 |
|          | PCR prime: Forward: acactccagctgggtgccaaagga                                     | Reverse: aactggtgtcgtggag                       |
| miR5538  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggtgctgtg              |                                                 |
|          | PCR prime: Forward: acactccagctgggtgccaaagga                                     | Reverse: aactggtgtcgtggag                       |
| miR4995  | stem-loop RT-primer: ctcaactggtgtcgtggagtccggcaattcagttgaggtgctgtg              |                                                 |
|          | PCR prime: Forward: acactccagctgggtgccaaagga                                     | Reverse: aactggtgtcgtggag                       |
| miR#     | Universal reverse primer: Reverse: aactggtgtcgtggag                             | Reverse: aactggtgtcgtggag                       |
| 5S rRNA  | PCR prime: Forward: tctcggctctccagggcgt                                       | Reverse: atctcggctctcaggt                         |

#: 12 identified miRNAs.

doi:10.1371/journal.pone.0123083.t001

### Table 2. Target gene primers used for qRT-PCR.

| Target gene | Sequence(5'-3')                  | PCR prime Forward/Reverse                      |
|-------------|----------------------------------|-------------------------------------------------|
| SPL 16      | Forward: AAGTGCTTGGTGTTATCCTCT   | Reverse: CCCTCCTCATTAG                          |
| SRPK4       | Forward: AACCTTCTCACACTCATCA      | Reverse: CCCTCCTCATTAG                          |
| WRKY19      | Forward: GCTGACCAATTGGAGTAGAAG    | Reverse: GTGAGTGAATTGGTAACTCT                   |
| GAMYB       | Forward: CATAGTGAGCTGACACTCCATTAG| Reverse: GTGAGTGAATTGGTAACTCT                   |
| DFR         | Forward: CCTAGTGAGCTGAGTGAGTAGA  | Reverse: GTGAGTGAATTGGTAACTCT                   |
| F-box       | Forward: TCTCCCTCCTCAAACTC       | Reverse: GTGAGTGAATTGGTAACTCT                   |

doi:10.1371/journal.pone.0123083.t002
in root, flower and fruit tissues were quantified relative to this value. The whole assay was repeated three times in order to ensure the reliability of the assay data. The standard deviations of the data were obtained from three independent experiments. The 6 relative targets expression levels were also quantified using the $2^{-\Delta\Delta C_t}$ method.

Results

Computational identification of banana miRNAs

The conservation of miRNA families among various plant species has facilitated the prediction of conserved miRNAs by EST and/or GSS analysis in various plants [36]. Using this method, we identified 32 conserved miRNAs derived from 16 EST and 16 GSS sequences, belonging to 13 miRNA families (Table 3). Eight and four predicted miRNA members belonged to the families of miR156/157 and miR166, respectively, while three members belonged to the family of miR159/319, miR160, and miR162, respectively. The remaining miRNA families had no more than two members; among them, 28 miRNAs (87.5%) were derived from *M. acuminata* (AA and AAA) and 4 miRNAs (miR159, miR397, miR399a*, and miR399a) were derived from *Musa* ABB group; miR4995 and miR5538 are the new families in banana that have not been reported previously (Table 3).

Though the lengths of miRNA precursors varied from 77 nt to 176 nt (Table 3), the lengths of mature miRNAs were mainly confined to 21 nucleotides (65.6% = 21/32), except five (15.6% = 5/32) miRNAs (miR156a-5p, miR156g, miR319c, miR162, and miR172b) and seven (21.9% = 7/32) miRNAs (miR156, miR156a-3p, miR156h, miR157b*, miR166, miR167d, and miR5538) with 20 and 22 nucleotides, respectively. The A+U contents of the pre-miRNAs ranged from 43.53% to 58.78%, with an average of 49.32%, which met the previous criteria. Mature miRNA sequences have been reported to be evenly located on the 2 arms of the stem-loop hairpin structures of potential pre-miRNAs [40]. In our study, slightly more mature miRNAs were located in the 3′ end of the pre-miRNA than were located in the 5′ end (17:15). In addition, 5 pairs of miRNA/miRNA* were located at the same precursors. Among the miRNAs, 65.6% (21/32) of members begin with a 5′ uracil. These results are in agreement with previous observations [19; 41,42]. All 33 miRNAs had fewer than six mismatched base pairs between the region of the mature miRNA sequences and its opposite arms. MFE is an important parameter in determining the stability of the perfect or near-perfect secondary hairpin structure of pre-miRNA, as a lower MFE value generally indicates a more stable secondary structure. In the present study, the average value of MFE was -45.3 kcal/mol with a range of -19.1 kcal/mol to -68.8 kcal/mol.

Conservation of the newly identified banana miRNAs

The miRNA families found in the miRBase are classified as conserved or unique depending on the number of species in which they have been identified. miRNA families registered in single species are considered to be unique, and those registered in two species and three or more species as moderately and highly conserved, respectively [43]. We analyzed 14 miRNAs registered in miRBase in 19 species. Multiple sequence alignment for 14 miRNAs registered in miRBase in 19 species have shown in S1 File. Based on our results (Fig 2), 11 of the 13 miRNA families are conserved within the vegetable kingdom, namely *Arabidopsis thalinana*, *Nicotiana tabacum*, *Malus domestica*, *Cucumis melo*, *Glycine max*, *Oryza sativa*, and *Zea mays*, but the miR4995 and miR5538 family are exceptions. miR4995 and miR5538 show conservation with that of *Glycine max* and *Oryza sativa* respectively.
Prediction of banana miRNA target genes

Based on the complementarity between miRNAs and their target genes in plants, all gene-coding sequences in the banana genome database (http://banana-genome.cirad.fr/download/musa_cds.fna.gz) were used to predict banana miRNA targets by psRNATarget (http://plantgrn.noble.org/psRNATarget/). 244 miRNA:target pairs between 32 miRNAs and 121 mRNAs have been predicted.

Table 4 shows some potential target-genes harboring binding sites for 13 conserved miRNA families. A complete list of predicted targets is depicted in S1 Table. Except for miR5538, every miRNA has more than one target. Further analysis found that most of the highly conserved
miRNA families target transcription factors, such as the SBP family targeted by miR156, TCP and MYB families targeted by miR159/319, ARF family targeted by miR160, NAC family targeted by miR164, HD-Zip family targeted by miR166, and WRKY family targeted by miR399 (Table 4). In addition, other predicted target genes are involved in diverse metabolic and physiological processes, such as dehydrogenase, decarboxylase, helicase, synthase, protein kinases, oxidoreductase, binding protein, lipase, transporters, disease resistance protein, f-box proteins, zinc finger protein, heat shock protein, and laccase (Table 4).

miRNAs and target confirmation by RT-qPCR

In order to examine the expression of these miRNAs, stem-loop RT-PCR were performed using a total RNA isolated from roots, leaves, flowers, and fruits. Using this approach, we were able to detect a positive signal of the expected size at 70nt (data not shown), and to confirm the presence of these miRNAs. 12 miRNAs from 11 families were chosen randomly in this experiment. All the PCR products were cloned and verified by sequencing. 12 potential miRNAs were expressed and exhibited different expression patterns in four tissues (Fig 3). The expression levels of miR164, HD-Zip family targeted by miR166, and WRKY family targeted by miR399 (Table 4). In addition, other predicted target genes are involved in diverse metabolic and physiological processes, such as dehydrogenase, decarboxylase, helicase, synthase, protein kinases, oxidoreductase, binding protein, lipase, transporters, disease resistance protein, f-box proteins, zinc finger protein, heat shock protein, and laccase (Table 4).

Fig 2. The statistical of the miRNAs distribution in different plants. Digital number on behalf of the members of the family. The species names are abbreviated as follows: Eu, Eudicotyledons; Mo, Monocotyledons; ath, Arabidopsis thaliana; ghr, Gossypium hirsutum; nta, Nicotiana tabacum; cme, Cucumis melo; cpa, Carica papaya; mdm, Malus domestica; csi, Citrus sinensis; sal, Salvia sclarea; ptc, Populus trichocarpa; acq, Aquilegia caerulea; hbr, Hevea brasiliensis; bcy, Bruguiera cylindrica; gma, Glycine max; vvi, Vitis vinifera; osa, Oryza sativa; sbl, Sorghum bicolor; sof, Saccharum officinarum; mpa, Musa paradisiaca; tae, Triticum aestivum; zma, Zea mays.

doi:10.1371/journal.pone.0123083.g002
level in fruits, roots, and flowers, respectively; thus, the conserved miRNAs were expressed in a tissue-specific manner.

| miRNA family | Target protein                                                                 | Function                           |
|--------------|--------------------------------------------------------------------------------|------------------------------------|
| miR156/157   | Squamosa promoter-binding-like protein                                         | Transcription factor               |
|              | OsSPL16—SBP-box gene family member                                             | Transcription factor               |
|              | ELMO/CED-12 family protein                                                     | Signal transduction                |
|              | CTP synthase 1                                                                | Metabolism                         |
|              | Putative Abscisic acid 8'-hydroxylase 3                                       | Metabolism/Signal transduction     |
|              | RNA recognition motif containing protein                                       | Metabolism                         |
| miR159/319   | Transcription factor GAMYB                                                    | Transcription factor               |
|              | TCP family transcription factor                                                | Transcription factor               |
|              | Chorismate mutase, chloroplastic                                               | Metabolism                         |
|              | Putative Nuclear polyadenylated RNA-binding protein 4                         | Transcription factor               |
|              | Putative Cysteine-rich receptor-like protein kinase 15                         | Signal transduction                |
| miR160       | Auxin response factor                                                          | Transcription factor               |
|              | pentatricopeptide repeat-containing protein At1g68930                          | Metabolism                         |
| miR162       | Anthrailate synthase component I-1, chloroplastic                              | Growth development                 |
| miR164       | NAC domain-containing protein 21/22 (NAC21/22/NAM/CUC2)                        | Transcription factor               |
|              | UDP-glucuronic acid decarboxylase 1                                            | Transcriptional regulation         |
|              | E3 ubiquitin-protein ligase BRE1-like 1                                        | Signal transduction                |
| miR166       | HD-ZIP III                                                                     | Stress responses                   |
|              | SRPK4                                                                          | Transcription factor               |
|              | Putative 2-oxoglutarate dehydrogenase, mitochondrial                          | Metabolism                         |
|              | pectinacetylerase domain containing protein, expressed                        | Metabolism                         |
|              | Putative U-box domain-containing protein 33                                   | Signal transduction/protein metabolism |
|              | GDSL esterase/lipase At5g33370                                                 | Metabolism                         |
| miR167       | Putative tRNA-specific adenosine deaminase 1                                   | Metabolism                         |
|              | Putative U-box domain-containing protein 52                                   | Signal transduction/protein metabolism |
| miR169       | Glutamate synthase [NADH], amyloplastic                                        | Metabolism                         |
| miR172       | amine oxidase-related                                                          | Metabolism                         |
| miR397       | Laccase                                                                        | Stress response                    |
|              | Putative pectinesterase/pectinesterase inhibitor 28                           | Metabolism                         |
|              | Putative S-(hydroxymethyl)glutathione dehydrogenase                            | Metabolism/oxidative stress response |
|              | Putative Heat shock 70 kDa protein 4L                                          | Stress response                    |
| miR399       | Probable WRKY transcription factor 19                                           | Transcription factor               |
|              | Putative Zinc finger protein NUTCRACKER                                        | Transcription factor               |
|              | 2-oxoglutarate/malate translocator, chloroplastic                              | Transport                          |
|              | DEAD-box ATP-dependent RNA helicase 37                                         | Metabolism                         |
| miR4995      | F-box protein At2g32560                                                         | Transcription factor               |
|              | zinc knuckle domain containing protein                                         | Metabolism                         |
| miR5538      | Dihydroflavonol-4-reductase                                                    | Metabolism                         |

Table 4. Putative target genes of miRNA candidates identified in banana.
In order to compare the expression of targets with that of the miRNAs, qRT-PCR was also used to confirm the expression of computationally predicted banana miRNA target genes in leaves, roots, flowers, and fruits; these miRNA target genes included SPL16, SRPK4, GAMYB, WRKY19, F-box, and DFR. The expression patterns of target genes in banana tissues were negatively correlated with the miRNA quantification results (Fig 4): the lowest expression levels of the miR156d, miR166b, miR399a, miR4995, and miR5538 target genes were consistent with the highest expression levels of miR156d, miR166b, miR399a, miR4995, and miR5538 in the flower tissue. The lowest expression level of the miR319 target gene was consistent with the highest expression level of miR319 in the fruit tissue. The higher expression levels of the six target genes were consistent with the lower expression of the six miRNAs in four tissues.

Discussion

Although miRNAs have been extensively studied for several years, no systematic study has been performed on the banana, one of the most important subtropical crops in the world. Recently, D’Hont et al (2012) and Davey et al (2013) identified 37 miRNA families and 42 miRNA families from Musa A- and B-genome, respectively, using bioinformatics analysis: those predicted miRNAs have not been confirmed by experiments for their precise sequences. Thus, banana miRNAs still remain largely unknown. Though there are a few computational approaches to identify miRNAs of plants, animals and even microorganisms, EST/GSS analysis has some advantages over other methods [19,44].

Using EST and GSS analysis, we identified 32 miRNAs belonging to 13 miRNA families. Among them, miR4995 and miR5538 are the newly found families in banana that have not previously been reported. The frequency of miRNA from ESTs and GSSs was 0.037% and 0.051%, respectively, which is little higher than the rate of 0.01% (1 in 10,000 ESTs) reported by Zhang et al (2006). 11 miRNA families identified in the study were the same as previously reported results [27,32], which have, to some extent, verified the utility of the computational predication of potential miRNAs. Several studies showed that many miRNAs are conserved across different plants [1]; however, some miRNAs are species-specific. For instance, miR5538 family is present in rice but not in Arabidopsis, suggesting that it might be restricted to monocots [45]. Cloning and sequencing validated 12 miRNAs belonging to 11 families; moreover, the ubiquitous expression of these 12 miRNAs could be detected in different banana tissues ranging from roots, leaves, flowers and fruits by qRT-PCR and some showed differential expression in different
tissues. Our qRT-PCR analysis validated the miRNA prediction in banana, and their preferential expression can provide important clues about the miRNAs' function.

The prediction of miRNA target genes is important for understanding miRNA gene-regulation. Currently, bioinformatics methods based on the high homology between miRNAs and target genes are considered one of the most efficient approaches used to verify target genes, and has been employed in numerous studies [22; 46–49]. Using the newly identified miRNAs in banana as query sequences to blast against the banana genome database, we were able to predict various miRNA target genes involved in transcription, translation, stress response, structural component, development, and metabolism in the banana (Table 4).

We predicted 244 miRNA:target pairs for 32 miRNAs giving an average miRNA to target ratio of 7.6 (244/32), which is greater than the ratios of 2.8 for the apple [23] and 4.1 for the castor bean [48]. Consistent with previous finding, most of these targets in the banana were plant-specific transcription factors, such as SBP, TCP, and MYB, NAC, ARF, and the HD-ZIP family, which were also predicted as targets of miR156/157, miR159/319, miR164, miR160 and miR166 in Arabidopsis [46,50], tobacco [42], citrus [22], apples [23], castor beans [48], and papayas [49]. MiR397 family targets Laccase were also conserved among plants; however, the

**Fig 4.** The relative expression level of selected six miRNAs (a) and their targets (b) in leaf, root, flower, and fruit of banana. The horizontal axis indicates the 6 miRNAs and their targets, and the vertical axis indicates the relative expression.

doi:10.1371/journal.pone.0123083.g004
other three miRNA families target genes that are found in a broad range of plants were not predicted to target genes with similar functions in the current banana database: miR162:DCL1 (Dicer1), miR167:ARF (auxin response factor), and miR172:AP2 (APETALA2).

qRT-PCR was used to confirm the expression of potential banana miRNAs and their targets. The expression patterns between miRNAs and their targets showed a negative correlation [51,52]. In our study, the lower expression levels of the six target genes were consistent with the higher expression of the six miRNAs in four tissues, indicating that they may be actively cleaved by miRNAs (Fig 4). Squamosa promoter-binding protein-like (SPL) is the target of miR156 family. SPL genes encode plant-specific transcription factors that play important roles in fruit ripening [53], ovary and fruit development [54], grain size, shape, and quality [55], plant phase transition, plant architecture, and gibberellins signaling [56]. Our study showed that the decrease of miR156d expression may lead to increased expression of the SPL16 in roots and fruits in the banana. MiR319 plays an important role in flower organ identity and cold-induced stress [51,57]. The lowest expression level of the miR319m target gene (GAMYB) was consistent with the highest expression level of miR319m in the banana fruits. GAMYB has been demonstrated to play an important role in flower development, especially in stamen and anther development [58–60]; furthermore, GAMYB is also involved in GA-mediated programmed cell death (PCD) in the aleurone and tapetum [61].

MiR5538 showed the highest expression in flowers, and the target of miR5538, DFR (Dihydroflavonol-4-reductase), one of the key enzymes involved in flavonoids biosynthetic pathway, was expressed at the lowest level in the banana flowers. Flavonoids are involved in several biological processes including pigmentation of flowers, seed development, pollen viability, auxin transport regulation, protection against UV-B injury, defense against pathogens, and pests [62–64]; this may indicate that miR5538 has an important role in flower development and pollen viability in banana. F-box proteins, as targets for miR393/394, are reported to be differentially regulated by stress conditions and to play important roles in the abiotic stress response pathway [10,65]. In our study, miR4995 targets F-box protein (At2g32560), zinc knuckle domain containing protein, and bifunctional aspartokinase/homoserine dehydrogenase 1. The function of miR4995 is still not clear. Further studies will need to be performed in order to determine the regulation pattern of banana miRNAs and their targets.

Supporting Information

S1 File. Multiple sequence alignment for 14 miRNAs registered in miRBase in19 species. (DOC)

S1 Table. Putative target genes of miRNA candidates identified in banana. (XLS)

Acknowledgments

This research was supported by grants from the Natural Science Foundation of Hainan Province (No.312050, 312051), the National Nonprofit Institute Research Grant of CATAS-ITBB (No. ITBB130201), the National Natural Science Foundation of China (No. 31201603), and the Major Technology Project of Hainan Province (No. ZDZX2013023-1). The authors thank Dr Zhang Jiaming (Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences) for the helpful advice in preparing the manuscript.
Author Contributions
Conceived and designed the experiments: JYW. Performed the experiments: JC HRS MYR. Analyzed the data: JYW JC. Contributed reagents/materials/analysis tools: JYW. Wrote the paper: JYW RJF. Supervised the work: YDZ.

References
1. Bartel DP. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell, 116:281–297. doi:http://dx.doi.org/10.1016/S0092-8674(04)00045-5 PMID: 14744438
2. Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS. (2005) Nuclear processing and export of microRNAs in Arabidopsis. Proc Natl Acad Sci USA, 102:3691–3696. doi: 10.1073/pnas.0405570102 PMID: 15738428
3. Papp I, Mette MF, Aufsatz W, Daxinger L, Schauer SE, Ray A, et al. (2003) Evidence for nuclear processing plant microRNA and short interfering precursors. Plant Physiol. 132:1382–1390. doi: http://dx.doi.org/10.1104/pp.103. PMID: 12857820
4. Khvorova A, Reynolds A, Jayasena SD. (2007) Functional siRNAs and miRNAs exhibit strand bias. Cell, 131: 41–49. doi:10.1016/S0092-8674(03)00801-8
5. Hake S. (2003) MicroRNAs: A role in plant development. Curr Biol, 13: R851–R852. doi:10.1016/j.cub.2003.10.021 PMID: 14588265
6. Valencia-Sanchez MA, Liu J D, Hannon G J, Parker R. (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. Gene Dev, 20: 515–524. doi:10.1101/gad.1399806 PMID: 16510870
7. Dong QH, Han J, Yu HP, Wang C, Zhao MZ, Liu H, et al. (2012) Computational identification of microRNAs in strawberry expressed sequence tags and validation of their precise sequences by miR-RACE. J. Heredity, 103(2):268–277. doi: 10.1093/jhered/esr127
8. Llave C, Kasschau KD, Rector MA, Carrington JC. (2002) Endogenous and silencing-associated small RNAs in plants. Plant Cell, 14: 1605–1619. doi: http://dx.doi.org/10.1105/tpc.003210 PMID: 12119378
9. Floyd SK, Bowman JL. (2004) Gene regulation: ancient microRNA target sequences in plants. Nature, 428:485–486. doi:10.1038/428485a PMID: 15057819
10. Zhang BH, Wang QL, Cobb GP, Anderson TA. (2005) Identification and characterization of new plant microRNAs using EST analysis. Cell Res, 15: 336–360. doi: 10.1038/sj.cr.7290302 PMID: 15916721
11. Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, Vance V, et al. (2005) Computational prediction of miRNAs in Arabidopsis thaliana. Genome Res, 15:78–91. doi:10.1101/gr.2908205 PMID: 15632092
12. Archak S, Nagaraju J (2007) Computational prediction of rice (Oryza sativa) miRNA targets. Genomics Proteomics Bioinformatics, 5:196–206. doi:10.1016/j.gpb.2007.03.007-8 PMID: 18267301
13. Zhang BH, Pan XP, Anderson TA. (2006) Identification of 188 conserved maize microRNAs and their targets. FEBS Lett., 580:3753–3762. doi: 10.1016/j.febslet.2006.05.063 PMID: 16780841
14. Colaiacovo M, Subacchi A, Bagnaresi P, Lamontanara A, Cattivelli L, Faccioli P, et al. (2010) A computational-based update on microRNAs and their targets in barley (Hordeum vulgare L.). BMC Genomics, 11:595–609. doi:10.1186/1471-2164-11-595 PMID: 20969764
15. Du JF, Wu YJ, Fang XF, Cao JX, Zhao L, Tao SH. (2010) Prediction of sorghum miRNAs and their targets with computational methods. Chin Sci Bull., 55:1263–1270. doi:10.1007/s11434-010-0035-4
16. Zhang BH, Wang QL, Wang KB, Pan XP, Liu F, Guo T, et al. (2007) Identification of cotton microRNAs and their targets. Gene, 397:26–37. doi:10.1016/j.gene.2007.03.020 PMID: 17574351
17. Zhang B, Pan X, Stellwag EJ. (2008) Identification of soybean microRNAs and their targets. Planta, 229:161–182. doi:10.1007/s00425-008-0818-x PMID: 18615805
18. Zhang WW, Luo YP, Gong X, Zeng WH, Li SG. (2009) Computational identification of 48 potato microRNAs and their targets. Comput Biol Chem, 33: 84–93. doi: 10.1016/j.compbiolchem.2008.07.006 PMID: 18723398
21. Yin Z, Li C, Han X, Shen F. (2008) Identification of conserved microRNAs and their target genes in tomato (Lycopersicon esculentum). Gene, 414:60–66. doi: 10.1016/j.gene.2008.02.007 PMID: 18387754

22. Song CN, Jia QD, Fang JG, Li F, Wang C, Zhang Z. (2010) Computational identification of citrus microRNAs and target analysis in citrus expressed sequence tags. Plant Biol (Stuttgart), 12:927–934. doi: 10.1111/j.1438-8677.2009.00300.x PMID: 21040308

23. Ye KY, Chen Y, Hu XW, Guo JC. (2013) Computational identification of microRNAs and their targets in apple. Genes Genom., 35:377–385. doi: 10.1007/s13258-013-0070-z

24. Zhou ZS, Huang SQ, Yang ZM. (2008) Bioinformatic identification and expression analysis of new microRNAs from Medicago truncatula. Biochem Biophys Res Commun, 374:538–42. doi: 10.1016/j.bbrc.2008.07.083 PMID: 18662674

25. Xie FL, Huang SQ, Guo K, Xiang AL, Zhu YY, Nie L, et al. (2007) Computational identification of novel microRNAs and targets in Brassica napus. FEBS Lett, 581:1464–1474. doi: 10.1016/j.febslet.2007.02.074 PMID: 17367786

26. Pani A, Mahapatra RK. (2013) Computational identification of microRNAs and their targets in Catharanthus roseus expressed sequence tags. Genomics Data, 1: 2–6. doi: 10.1016/j.gdata.2013.06.001 PMID: 23311594

27. D’Hont A, Denoue F, Aury J-M, Baurens F-C, Carreel F, Garsmeur O, et al. (2012) The banana (Musa acuminata) genome and the evolution of monocotyledonous plants. Nature 488: 213–217. doi: 10.1038/nature11241 PMID: 22801500

28. Baurens FC, Bocs S, Rouard M, Matsumoto T, Miller RN, Rodier-Goud M, et al. (2010) Mechanisms of haplotype divergence at the RGA08 nucleotide-binding leucine-rich repeat gene locus in wild banana (Musa balbisiana). BMC Plant Biol., 10:149. doi: 10.1186/1471-2229-10-149 PMID: 20637079

29. Simmonds N. W. (1962) The Evolution of the Bananas. Longman press. 101p

30. FAOSTAT, banana production. http://faostat3.fao.org/faostat-gateway/go/to/browse/Q/*

31. Kozomara A, Griffiths-Jones S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucl. Acids Res. 42:D68–D73. doi: 10.1093/nar/gkt1181 PMID: 24275496

32. Davey MW, Gudimella R, Harikrishna JA, Sin LW, Khalid N, Keulemans J. (2013) A simple and efficient method for isolating RNA from pine trees. J. Plant Biochem. Biotech., 22:118–123. doi: 10.1111/j.1365-313X.2012.01354.x PMID: 24094114

33. Zhou ZS, Huang SQ, Yang ZM. (2008) Bioinformatic identification and expression analysis of new microRNAs from Medicago truncatula. Biochem Biophys Res Commun, 374:538–42. doi: 10.1016/j.bbrc.2008.07.083 PMID: 18662674

34. Dai XB, Zhao PX. (2011) psRNATarget: A Plant Small RNA Target Analysis Server. Nucl. Acids Res., 39: w155. doi: 10.1093/nar/gkr1181 PMID: 24275496

35. Livak KJ, Schmittgen TD. (2001) Analysis of Relative Gene Expression Data Using Real Time Quantitative PCR and the 2−ΔΔCT Method. Methods, 25, 402–408. doi: 10.1006/meth.2001.1262 PMID: 11846609

36. Gorodkin J, Høvgaard JH, Enstroer M, Sawera M, Jensen P, Ohman M, et al. (2006). MicroRNA sequence motifs reveal asymmetry between the stem arms. Comput. Biol. Chem., 30: 249–254. doi: 10.1016/j.compbiolchem.2006.04.006 PMID: 16798093

37. Frazier TP, Xie FL, LFrestadler A, Burklew CE, Zhang BH. (2010) Identification and characterization of microRNAs and their target genes in tobacco (Nicotiana tabacum). Planta, 232(6):1289–1308. doi: 10.1007/s00425-010-1255-1 PMID: 20803216

38. Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA. (2006). Evidence that miRNAs are different from other RNAs. Cell. Mol. Life Sci. 63, 246–254. doi: 10.1007/s00018-005-5467-7 PMID: 16395542

39. Dryanova A, Zakharov A, Gulicuka PJ. (2008) Data mining for miRNAs and their targets in the Triticaceae. Genome, 51(6): 433–443. doi: 10.1139/G08-025 PMID: 18521122

40. Frazier TP, Sun G, Burklew CE, Zhang BH. (2011) Salt and drought stresses induce the aberrant expression of microRNA genes in tobacco. Mol. Biotechnol., 49 (2):159–165. doi: 10.1007/s12033-011-9387-5 PMID: 21359858
45. Wei LQ, Yan LF, Wang T. (2011) Deep sequencing on genome-wide scale reveals the unique composition and expression patterns of microRNAs in developing pollen of Oryza sativa. Genome Biol., 12(6): R53–R68. doi: 10.1186/gb-2011-12-6-r53 PMID: 21679406

46. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. (2002) Prediction of plant microRNA targets. Cell 110:513–520. doi: 10.1016/S0092-8674(02)00863-2 PMID: 12202040

47. Bonnet E, Wuyts J, Rouze P, Peer VY. (2004) Detection of 91 potential in plant conserved plant micro-RNAs in Arabidopsis thaliana and Oryza sativa identifies important target genes. Proc Natl Acad Sci USA, 101:11511–11516. doi: 10.1073/pnas.0404025101 PMID: 15272084

48. Zeng CY, Wang WQ, Zheng Y, Chen X, Bo WP, Song S, et al. (2009) Conservation and divergence of microRNAs and their functions in Euphorbiaceous plants. Nucl. Acids Res., 38(3):981–995. doi: 10.1093/nar/gkp1035 PMID: 19942686

49. Liang G, Li Y, He H, Wang F, Yu DQ. (2013) Identification of miRNAs and miRNA-mediated regulatory pathways in Carica papaya. Planta, 238:739–752. doi: 10.1007/s00425-013-1929-6 PMID: 23851604

50. Sunkar R, Zhu J K. (2004) Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. Plant Cell, 16(8): 2001–2019. doi: http://dx.doi.org/10.1105/tpc.104 PMID: 15258262

51. Thiebaut F, Rojas CA, Almeida KL, Grativol C, Domiciano GC, Lamb CR, et al. (2012) Regulation of miR319 during cold stress in sugarcane. Plant Cell Environ., 35 (3):502–512. doi: 10.1111/j.1365-3040.2011.02430.x PMID: 22017483

52. Unver T, Parmaksiz I, Dündar E. (2010) Identification of conserved micro-RNAs and their target transcripts in opium poppy (Papaver somniferum L.). Plant Cell Rep., 29 (7):757–769. doi: 10.1007/s00299-010-0862-4 PMID: 20443006

53. Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, et al. (2006) A naturally occurring epigallocatechin gallate inhibits the germination of rice seeds by targeting the rice homeobox gene OsSPL16. Planta, 223(5): 604–618. doi: 10.1007/s00425-005-0618-6.

54. Wang SK, Wu K, Yuan QB, Liu XY, Liu ZB, Lin X, et al. (2012) Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet., 44(8):950–954. doi: 10.1038/ng.2327 PMID: 22729225

55. Chen XB, Zhang ZL, Liu DM, Zhang K, Li AL, Mao L. (2010) SQUAMOSA promoter-binding protein-like genes, GAMYB-like genes, and MYB65 are microRNA-regulated genes that redundantly facilitate anther development. Plant Cell, 17:705–721. doi: http://dx.doi.org/10.1105/tpc.104 PMID: 15487003

56. Wang SK, Wu K, Yuan QB, Liu XY, Liu ZB, Lin X, et al. (2012) Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet., 44(8):950–954. doi: 10.1038/ng.2327 PMID: 22729225

57. Schwab R, Palatnik JF, Riester M, Schmid M, Weigel D. (2005) Specific effects of microRNAs on the plant transcriptome. Dev. Cell, 8, 517–527. doi: 10.1016/j.devcel.2005.01.018 PMID: 15809034

58. Murray F, Kalla R, Jacobsen J, Gubler F. (2003) A role for HvGAMYB in anther development. Plant J., 33: 481–491. doi: 10.1046/j.1365-313X.2003.01641.x PMID: 12581306

59. Liu Z, Bao W, Liang W, Yin J, Zhang D. (2010) Identification of gamyb-4 and analysis of the regulatory role of GAMYB in rice anther development. J Integr Plant Biol., 52(6): 670–678. doi: 10.1111/j.1744-7909.2010.00959.x PMID: 20590996

60. Millar AA, Gubler F. (2005) The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. Plant Cell, 17:705–721. doi: http://dx.doi.org/10.1105/tpc.104 PMID: 15722475

61. Aya K, Ueguchi-Tanaka M, Kondo M, Hamada K, Yano K, Nishimura M, et al. (2009) Gibberellin modulates anther development in rice via the transcriptional regulation of GAMYB. Plant Cell, 21, 1453–1472. doi: http://dx.doi.org/10.1105/tpc.104 PMID: 19454733

62. Winkel-Shirley B (2001) It takes a garden. How work on diverse plant species has contributed to an understanding of flavonoid metabolism. Plant Physiol., 127:1399–1404. doi: http://dx.doi.org/10.1104/pp.1010675 PMID: 11743081

63. Martens S, Knott J, Seitz CA, Janvari L, Yub SN, Forkmann G. (2003) Impact of biochemical pre-studies on specific metabolic engineering strategies of flavonoid biosynthesis in plant tissues. Biochem Eng J, 14: 227–235. doi: 10.1016/S1369-703X(02)00224-3

64. Peer WA, Murphy AS (2007) Flavonoids and auxin transport: modulators or regulators? Trends in plant science, 12: 556–563. doi: 10.1016/j.tplants.2007.10.003 PMID: 18198522

65. Navarro L, Dunoyer P, Jay F, Arnold B, Dharmanis N, Estelle M, et al. (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science, 312:436–439. doi: 10.1126/science.1126088 PMID: 16627744