

Comparative Analysis Reveals Dynamic Changes in miRNAs and Their Targets and Expression during Somatic Embryogenesis in Longan (*Dimocarpus longan* Lour.)

Yuling Lin, Zhongxiong Lai*

Institute of Horticultural Biotechnology, Fujian Agriculture and Forestry, Fuzhou, Fujian, China

Abstract

Somatic embryogenesis (SE), which resembles zygotic embryogenesis, is an essential component of the process of plant cell differentiation and embryo development. Although microRNAs (miRNAs) are important regulators of many plant developmental processes, their roles in SE have not been thoroughly investigated. In this study, we used deep-sequencing, computational, and qPCR methods to identify, profile, and describe conserved and novel miRNAs involved in longan (*Dimocarpus longan*) SE. A total of 643 conserved and 29 novel miRNAs (including star strands) from more than 169 miRNA families were identified in longan embryogenic tissue using Solexa sequencing. By combining computational and degradome sequencing approaches, we were able to predict 2063 targets of 272 miRNAs and verify 862 targets of 181 miRNAs. Target annotation revealed that the putative targets were involved in a broad variety of biological processes, including plant metabolism, signal transduction, and stimulus response. Analysis of stage- and tissue-specific expressions of 20 conserved and 4 novel miRNAs indicated their possible roles in longan SE. These miRNAs were *dlo-miR156* family members and *dlo-miR166c** associated with early embryonic culture developmental stages; *dlo-miR26*, *dlo-miR160a*, and families *dlo-miR159*, *dlo-miR390*, and *dlo-miR398b* related to heart-shaped and torpedo-shaped embryo formation; *dlo-miR4a*, *dlo-miR24*, *dlo-miR167a*, *dlo-miR168a**, *dlo-miR397a*, *dlo-miR398b.1*, *dlo-miR398b.2*, *dlo-miR808* and *dlo-miR5077* involved in cotyledonary embryonic development; and *dlo-miR17* and *dlo-miR2089*-1* that have regulatory roles during longan SE. In addition, *dlo-miR167a*, *dlo-miR808*, and *dlo-miR5077* may be required for mature embryo formation. This study is the first reported investigation of longan SE involving large-scale cloning, characterization, and expression profiling of miRNAs and their targets. The reported results contribute to our knowledge of somatic embryo miRNAs and provide insights into miRNA biogenesis and expression in plant somatic embryo development.

Citation: Lin Y, Lai Z (2013) Comparative Analysis Reveals Dynamic Changes in miRNAs and Their Targets and Expression during Somatic Embryogenesis in Longan (*Dimocarpus longan* Lour.). PLoS ONE 8(4): e60337. doi:10.1371/journal.pone.0060337

Editor: Meng-xiang Sun, Wuhan University, China

Received: November 8, 2012; **Accepted:** February 25, 2013; **Published:** April 11, 2013

Copyright: © 2013 Lin, Lai. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by the National Natural Science Foundation of China (31078717, 31272149, and 31201614), the Doctoral Program of Higher Education of the Chinese Ministry of Education (20093515110005 and 20123515120008), and Fujian Provincial Natural Science Foundation of China (2012J05042). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: laizx01@163.com

Introduction

Longan (*Dimocarpus longan* Lour.) is a tropical/subtropical fruit tree in the family Sapindaceae. In this species, embryo development status strongly influences seed size, fruit quality, fruit set, and yield. To improve longan fruit quality and yield, researchers have consequently focused much attention on mechanisms of longan embryonic development. Advances have been limited, however, primarily because of the extreme genetic heterozygosity exhibited by this species and the difficulty of sampling the early embryos. Although the longan somatic embryogenesis (SE) system has been widely used as a model system for investigating *in vitro* and *in vivo* regulation of embryogenesis in woody plants, most studies of longan SE have focused on a few selected aspects related to cell biology, molecular biology, and proteomics [1–5]. There are no reported studies devoted to identification and expression of microRNAs (miRNAs)—an important group of plant regulators—during longan SE.

miRNAs, a class of single-stranded noncoding RNAs that are 18–25 nt in length, negatively regulate gene expression by cleaving targeted mRNAs and repressing translation [6]. There is increasing evidence that plant miRNAs play important roles in various processes, including leaf morphogenesis and polarization, lateral root formation, floral induction, meristem boundary formation, floral organ identity, and reproduction [7–10]. SE, which resembles zygotic embryogenesis, is strongly associated with plant cell differentiation and embryo development—a process during which many specific miRNAs are undoubtedly expressed [11,12]. Several studies have accordingly shown that miRNA-mediated repression of target transcripts plays a role in embryogenesis [13,14] and that certain miRNAs may function at specific developmental stages [11,15–18]. In *Arabidopsis*, miRNAs enable pattern formation during embryogenesis [14] and regulate the timing of embryo maturation by repressing several key targets [13]. *AGO1*, which is involved in miRNA function, is specifically expressed during carrot SE [19], and *ptAGO9L* as well as miRNAs

such as *Pta-miR166* and *Pta-miR167* exhibit spatiotemporally variable expression in loblolly pine zygotic embryos and female gametophytes [18]. These observations suggest that a miRNA expression controlling system is required for SE in these species. The regulatory role of miRNAs during plant embryogenesis is still not well understood, however, and few relevant studies have been conducted in non-model plants. No studies have attempted to detect miRNAs in longan.

In this study, we used Solexa sequencing and an integrated bioinformatics analysis approach to clone and identify conserved and novel miRNAs in longan. We then predicted, identified, and verified longan miRNA targets using prediction analysis, degradome sequencing, and modified RNA ligase-mediated amplification of cDNA ends (RLM-RACE). Finally, we comparatively analyzed expression patterns to reveal miRNA participation in longan SE. The results of this study provide new information regarding miRNA regulatory networks and contribute to our knowledge of the role of miRNAs in longan SE.

Results

Categories and Size Distribution of Small RNA Populations in *D. longan*

To identify miRNAs involved in longan SE, a pooled sRNA library was generated from embryogenic cultures and subjected to Solexa sequencing. We obtained 12,554,858 raw reads, comprising 447,851 low-quality and 12,107,007 high-quality reads. After removing low-quality reads and adapter sequences, the remaining 11,645,841 clean reads, which were 18–30 nt long, included 6,553,782 unique sequences. Only 378,143 (5.77%) of the unique sRNA sequences could be mapped to the longan transcriptome (Table 1). Most of the sequences (94.23%) could not be mapped.

The unique sRNAs were then compared against all plant miRNA precursors and mature miRNAs listed in miRBase (miRBase17); 18,299 (0.28%) of the unique sequences were found to be similar to known miRNAs. By performing a BLASTN search against the Rfam database, we then identified small RNAs corresponding to rRNA (0.49%), small nuclear RNA (snRNA; 0.03%), small nucleolar RNA (snoRNA; 0.03%), and tRNA (0.09%) in our unique sRNA dataset (Table 1). Most sRNA unique sequences (6,351,648; 96.92%) could not be annotated, however, which is consistent with results obtained in other plant studies [20–23]. The high percentage of unannotated sequences may be due to the limited number of species-specific genomes or ESTs in these databases, such that many potential miRNAs were missing; alternatively, plant sRNAs have not yet been adequately surveyed,

leading to incomplete data. The distribution of sRNAs detected in longan is given in Table 1.

The size distribution pattern of unique sRNA sequences generated by Solexa sequencing is summarized in Fig. 1. Distributions were uneven. Unique sRNAs were 18–25 nt long, with 24 nt lengths predominating (55.02%), followed by 22 nt (10.67%) and 20 nt (9.90%). The percentage of 24-nt sRNAs was much higher than that of 21 nt sRNAs. This result is consistent with that observed for most angiosperms [21,23,24] other than tomato [25], wheat [26], and *Populus* [27], in which the 21-nt length is more prevalent. The 21-nt to 24-nt length ratio is highly variable among plants, indicating the existence of significant differences in sRNA biogenesis pathways among different species [28].

Identification of Conserved miRNAs in *D. longan*

To identify conserved miRNAs in *D. longan*, unique sRNAs from longan were compared against mature plant miRNAs listed in miRBase17. As a result, 643 miRNAs belonging to more than 169 miRNA families were identified. These miRNAs varied in length from 18 nt to 25 nt. Most miRNAs were 24-nt long (34.07%), with 21 nt (20.2%), 20 nt (12.36%), and 19 nt (10.62%) lengths also heavily represented. A few were 25 nt long, such as *dlo-miR2095* and *dlo-miR3436**. After miRNAs with expression levels too low to be analyzed were removed, most miRNA families were represented by only one member. Some miRNA families were represented by multiple sRNAs, including *miR156* (17 members), *miR166* (12 members), *miR159* (8 members), *miR171* (9 members), *miR399* (5 members), *miR441* (6 members), and *miR1520* (6 members); these required further validation, however. Information about known miRNA families detected in longan, including number of members, sizes, and expression, is included in Table S1 in Supporting Information S1.

In addition to mature miRNAs, some miRNAs* with high expression were sequenced, including *dlo-miR408** (3727 reads), *dlo-miR946a** (1114 reads), *dlo-miR390a** (723 reads), *dlo-miR3441** (433 reads), *dlo-miR168a** (404 reads), *dlo-miR166c** (346 reads), and *dlo-miR1162** (300 reads) (Table S1 in Supporting Information S1). Compared with mature miRNAs, however, lower expression levels were observed during longan SE for most of the miRNAs*. This asymmetrical accumulation of transcripts may be due to preferential loading of the miRNA strand into the silencing complex, where it is protected from degradation; this is in contrast to the miRNA* strand, which is preferentially excluded from the silencing complex and subject to degradation [29].

Most of the miRNA families identified in longan are highly conserved across many plant species. For example, families *dlo-miR156* and *dlo-miR157* were a perfect match to families in *Arabidopsis thaliana*, *Zea mays*, *Vitis vinifera*, and *Populus*, and families *dlo-miR158–dlo-miR172* and *dlo-miR393–dlo-miR399* were highly homologous to families from other species. Dicot-specific miRNA families, including *miR158*, *miR161*, *miR163*, and *miR403* [30,31], were also detected in longan. In addition to conserved miRNAs, 21 non-conserved miRNA families were present in low abundance (2–123 reads) (Table S1 in Supporting Information S1); a few, including *dlo-miR535a* (9232 reads), *dlo-miR894* (6039 reads), *dlo-miR529* (881 reads) and *dlo-miR827* (753 reads), were expressed at high levels. Interestingly, several miRNA families considered to be monocot-specific—*miR437*, *miR443*, *miR444*, *miR445*, *miR528*, *miR1122*, and *miR1858*—were also found in our library (2–173 reads), suggesting that they also exist in dicots.

Table 1. Distribution of small RNAs among different categories in *Dimocarpus longan*.

category	Unique sRNA	Percent(%)	Total sRNA	Percent(%)
miRNA	18,299	0.28	1,208,973	10.38
rRNA	32,362	0.49	342,850	2.94
siRNA	–	–	–	–
snRNA	1,759	0.03	6,220	0.05
snoRNA	1,872	0.03	6,820	0.06
tRNA	5,633	0.09	158,587	1.36
unannotated	6,493,857	99.09	9,922,391	85.20
Total	6,553,782	100	11,645,841	100

doi:10.1371/journal.pone.0060337.t001

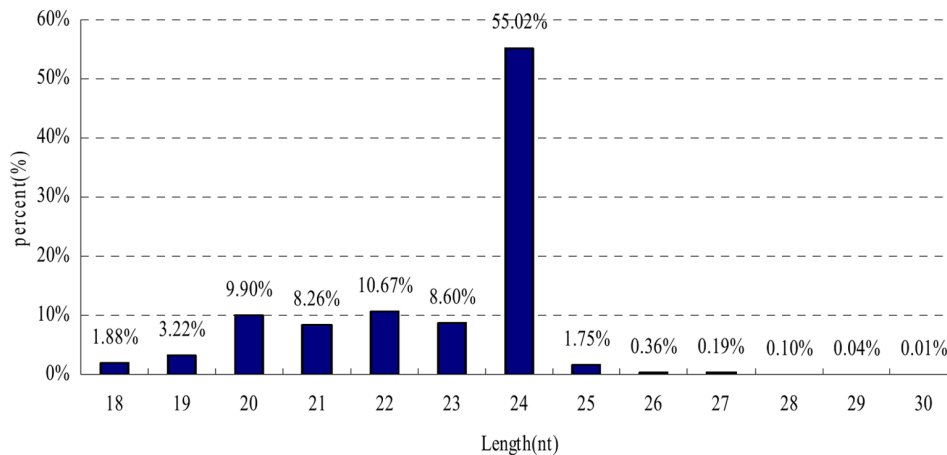


Figure 1. Length distribution of unique sequences in longan.
doi:10.1371/journal.pone.0060337.g001

Prediction of Potentially Novel Longan miRNAs via Longan Transcriptome and *Populus* Genome Analysis

In addition to conserved miRNAs, 29 potentially novel longan miRNAs were predicted from the remaining unannotated sRNA unique sequences—22 based on the longan transcriptome and 7 based on a *Populus trichocarpa* genome (Table 2). With the exception of *dlo-miR4* and *dlo-miR27*, which were represented by 2 genomic loci each, each novel longan miRNA was associated with a single genomic locus. The lengths of these miRNAs varied from 21 nt to 23 nt; 48.28% were 21 nt long. Based on Mfold calculations, their precursors had negative folding free energies ranging from -18.20 kcal/mol to -65.24 kcal/mol; the average was -40.03 kcal/mol, which was lower than the folding free energies reported for rice (-71.0 kcal/mol) and *Arabidopsis* (-59.5 kcal/mol) [26]. The number of reads obtained for these novel longan miRNAs varied from 5 to 8964; for example, *dlo-miR7* (8964 reads), *dlo-miR17* (1309), *dlo-miR9* (1144 reads), and *dlo-miR14* (609 reads) were highly expressed, while *dlo-miR1* (5 reads) and *dlo-miR18* (6 reads) were expressed at low levels. When nucleotide bias was analyzed, the nucleotide T (58.62%) was most frequent, followed by A (20.69%), C (10.34%), and G (13.79%).

To determine whether these novel miRNAs are conserved across other plant species, their sequences were compared with miRNA sequences of other organisms present in miRBase (miRBase19.0). This analysis revealed that *dlo-miR14*, *dlo-miR25*, and *dlo-miR26* were orthologous to *hbr-miR398*, *ptc-miR6478* [32], and *mtr-miR 4414b*, respectively, suggesting that these 3 miRNAs, newly identified in longan, are conserved in other plants. In addition to these 3 miRNAs, 11 other potentially novel miRNAs (4 based on longan and 7 based on *P. trichocarpa*) also had matches in the miRBase database. Although very similar to *zma-miR156d**, *gma-miR159a*, *aly-miR 162a**, *vun-miR164*, *zma-miR166c**, *aly-miR168a**, *aly-miR171c**, *aly-miR390a**, *gma-miR390a*, and *zma-miR398a**, they were not identical, which raises the question as to whether or not these miRNAs should be classified as known miRNAs. This problem is not unique: Pantaleo et al. [33] found that *miRC9* (a new miRNA) was very similar to *aqc-miR477e*, differing by only 2 mismatches. They classified *miRC9* as a new miRNA not belonging to the *miR477* family, however, because there were 6 mismatches between *miRC9* and *vi-miR477* and 7 mismatches between *miRC9* and *ptc-miR477*. In our study, the 11 potentially novel miRNAs in longan showed high homology

to known conserved miRNAs, and were consequently classified as known miRNAs rather than truly novel ones. As miRNAs are sequenced in an increasing number of species, this situation is expected to arise more frequently, leading to difficulties in miRNA classification.

Furthermore, miRNAs *dlo-miR22* through *dlo-miR27*, which were predicted based on the *P. trichocarpa* genome, cannot be accurately classified as longan-specific miRNAs because they may also be expressed in *P. trichocarpa*. If possible, species-specific miRNA prediction for non-model plants should rely on ESTs or GSSs originating from the studied species. When reference sequences originate from closely-related species, any predicted “novel” miRNAs will not be species-unique, but will instead be conserved across the selected species.

Prediction and Classification of Targets of Conserved and Non-conserved miRNAs in *D. longan*

To better understand the biological functions of longan miRNAs identified in our study, 2040 targets of 260 conserved miRNAs and 23 targets of 12 novel miRNA families were predicted using previously-described methods (Table S2 in Supporting Information S1).

Homologs of known miRNA targets were predicted for several conserved longan miRNAs, including *SPL10* as a target of *dlo-miR156*, *MYB33* for *dlo-miR159/319*, *NAC* for *dlo-miR164*, *ARF3/ARF8* for *dlo-miR167*, *SCL6* for *dlo-miR171*, *AP2* for *dlo-miR172*, *TIR1* for *dlo-miR393*, *F-box/NF-YC11* for *dlo-miR394*, and *laccase* for *dlo-miR397/408*. In addition, some miRNAs regulated more than one target. For example, while 23.85% (62) of the conserved miRNAs regulated only one target, 103 (39.62%) regulated 2–5 targets, and 19, including *miR2628* (169 targets), *miR2673* (118 targets), *miR854* (119 targets), *miR780* (51 targets), and *miR1442* (48 targets), had at least 20 targets; this suggests that these miRNAs are probably involved in a variety of biological processes during longan SE. Finally, some targets were regulated by multiple miRNAs. Examples include Unigene 62142 (ubiquitin-associated/T-S-N domain-containing protein) targeted by both *miR396* and *miR1520*, and Unigene 68246 (GRAS family transcription factor) regulated by *miR171*, *miR952*, *miR1081*, and *miR5078*. The above-predicted targets still required further experimental validation, however.

Based on Gene Ontology (GO) annotations, these potential miRNA targets appear to be involved in a broad variety of

Table 2. Longan-specific miRNAs identified from *Dimocarpus longan* transcriptome and populus genome.

miRNA	location	Sequence(5'–3')	Length (nt)	count	Mfe (kcal/mol)	5'/3'
<i>dlo-miR1</i>	Contig325306_Longan-RNA:27:90:+	TATGTATATATATGTGTAGAT	21	5	–18.2	3
<i>dlo-miR2</i>	Contig401887_Longan-RNA:25:90:–	TTGCTGGTGATGTGGTGGGGTG	22	9	–25.5	5
<i>dlo-miR3</i>	Contig424091_Longan-RNA:14:90:–	AAAGATCTTTAGGTTTTCTGT	20	15	–31.6	3
<i>dlo-miR4a</i>	Contig569989_Longan-RNA:1:81:–	ACGGGTTCAAAGGTTGACAGA	21	11	–27.2	3
<i>dlo-miR4b</i>	Unigene4793_Longan-RNA:488:584:+	ACGGGTTCAAAGGTTGACAGA	21	13	–32	3
<i>dlo-miR5</i>	Contig629412_Longan-RNA:1:95:–	TTGAGGAGTGGAAAGTCCAGAA	21	17	–19.75	3
<i>dlo-miR6</i>	Scaffold39496_Longan-RNA:47:139:–	TCAAGGGATGAAGATTTTAACT	22	213	–21	3
<i>dlo-miR7</i>	Unigene10300_Longan-RNA:116:231:–	TCATCGAACACAAGCCCATGC	22	8964	–47.5	3
<i>dlo-miR8</i>	Unigene12720_Longan-RNA:16:197:+	TTGATCAAATGTCCAAGGCTCA	22	611	–61.5	5
<i>dlo-miR9</i>	Unigene13717_Longan-RNA:777:907:–	AGTGAATGATGCGGGAGACAAAT	23	1144	–42.61	3
<i>dlo-miR10</i>	Unigene14204_Longan-RNA:50:162:+	TTCGAGCGCAAATTAATAGGA	21	12	–51.5	3
<i>dlo-miR11</i>	Unigene19708_Longan-RNA:14:99:–	TGATGCTGTAGATGATTCGGA	21	187	–20.7	5
<i>dlo-miR12</i>	Unigene3012_Longan-RNA:21:128:+	TGAAAATGAACTTAAGAGTTGC	22	120	–35.4	3
<i>dlo-miR13</i>	Unigene30135_Longan-RNA:22:153:–	AGAAGTTTTGATCTCGTAAGACA	23	164	–54	5
<i>dlo-miR14</i>	Unigene31710_Longan-RNA:49:145:–	GGAGCGACCTGAGATCACATG	21	609	–49.8	5
<i>dlo-miR15</i>	Unigene40670_Longan-RNA:84:181:+	TGGCTTTAATGAAGACTGGG	21	10	–37.89	3
<i>dlo-miR16</i>	Unigene47226_Longan-RNA:9:101:–	GGTATGGAAGGATTGGGTGCA	21	57	–35.2	5
<i>dlo-miR17</i>	Unigene54607_Longan-RNA:85:298:+	TGCAGCACAAAATACAGTCTGG	22	1309	–77.6	3
<i>dlo-miR18</i>	Unigene56582_Longan-RNA:219:296:+	TGGAAATTTGATGACTTACAT	22	6	–21.2	3
<i>dlo-miR19</i>	Unigene60375_Longan-RNA:59:202:–	CGTCATTCTAGTCGGATCATT	21	490	–55.81	5
<i>dlo-miR20</i>	Unigene67206_Longan-RNA:105:270:–	TTGAGGAAGTGAGCAAGAAAT	21	21	–65.24	5
<i>dlo-miR21</i>	Unigene8816_Longan-RNA:7:111:–	GAGGATTGATGGTAGACCTAAG	22	243	–22.8	3
<i>dlo-miR22</i>	LG_I:30129469:30129680:+	AAGTTTAAGAGGGGGTGTGAA	22	10	–41.5	5
<i>dlo-miR23</i>	LG_IV:14080863:14081082:+	GCTGGAGTAGCTCAGTTGGTT	21	130	–44.8	5
<i>dlo-miR24</i>	LG_V:15276189:15276277:–	TGAATGATTTCCGACCAGGCT	21	122	–40.8	3
<i>dlo-miR25</i>	LG_VIII:10042573:10042866:–	CCGACCTTAGCTCAGTTGGCAGA	23	72	–48.18	3
<i>dlo-miR26</i>	LG_XIII:9782450:9782561:+	TGTGAATGATGCGGGAGATAA	21	14	–40	3
<i>dlo-miR27a</i>	LG_II:21395857:21395930:–	TGGCGCTGCCGGAGTGGTTATC	22	10	–28.6	5
<i>dlo-miR27b</i>	LG_VIII:8501904:8501977:+	TGGCGCTGCCGGAGTGGTTATC	22	10	–32	5

doi:10.1371/journal.pone.0060337.t002

biological processes. Their functions could be divided into 5 categories: (1) plant metabolism, including biosynthesis of secondary metabolism, lignin catabolism, ethanol metabolism, and cell metabolism, (2) signal transduction and apoptosis, including calcium, B-cell receptor, TGF- β , and chemokine signaling pathways, (3) biological and abiotic stress processes (as for the NBS-LRR-type disease resistance protein, a target of *miR482*), (4) developmental processes (the case for *LEC* [Unigene 60033, a target of *miR476*] closely related to embryonic development), and (5) for many longan miRNA targets, genetic information processing, including categories such as spliceosome, ribosome, and RNA degradation, nucleotide excision, mismatch, and base excision repair, and homologous recombination. In addition, predicted targets not only included transcription factor genes (e.g., *AP2*, *NF- κ B*, *SPL*, *MYB*, and *WRKY*), but also novel genes of unknown function. The miRNAs identified in this study may control longan somatic embryo development by regulating the expression of these targets.

Identification of miRNA Cleavage Targets in *D.longan* Using Degradome Analysis

Bioinformatics analysis alone could not be used to determine whether the predicted miRNA targets were real or false positives. To verify predicted targets of known and novel miRNAs in the longan transcriptome, a high-throughput experimental approach—degradome sequencing—was therefore employed. More than 30 million raw reads were produced by Solexa sequencing. After processing, 72.14% of the resulting clean reads were 20- or 21-nt long; 58.81% (1,198,119) of the unique reads mapped perfectly to the longan transcriptome, indicating the high quality of the degradome library. The reads that mapped to the longan transcriptome were subjected to further analysis.

A total of 843 targets of 178 conserved miRNAs and 19 targets of 3 novel miRNAs (*dlo-miR1*, *dlo-miR2*, and *dlo-miR5*) were identified (Table S3 in Supporting Information S1). Similar to predictions based on sRNA library sequencing, degradome analysis indicated that most of the known miRNAs were capable of regulating multiple targets. These miRNAs included *dlo-miR854* (90 targets), *dlo-miR2673* (37 targets), *dlo-miR1046* (21 targets), *dlo-miR780* (13 targets), *dlo-miR1442* (10 targets), and *dlo-miR2628* (9

targets). The number of targets identified through degradome sequencing was less than that obtained from bioinformatics prediction methods, perhaps because the latter approach could not distinguish genuine targets from false ones. In contrast to known miRNAs, novel miRNAs (i.e., *dlo-miR1*, *dlo-miR2*, and *dlo-miR5*) appeared to have only a limited number of targets.

In longan, cleavage sites for most miRNAs were usually located in the CDS of their targets. Most targeted mRNAs harbored multiple cleavage sites, a phenomenon also observed in *Arabidopsis* and rice [34,35]. For example, the C2H2 (Unigene 11990) mRNA (targeted by *miR157*) was cleaved at nucleotide positions 211, 330, 331, 387, 404, and 1143. Only the site at position 211 was mediated by *miR157*, however. The remaining cleavage sites were split by other sRNAs or by unknown miRNAs, suggesting a complex degradation mechanism for these targets in plants. The cleavage sites in a given transcript may be targeted either by different miRNAs or the same one; the latter is the case with Unigene 67649 (similar to At4g28230), which has 7 cleavage sites (between positions 1411 and 1417) that are all split by *dlo-miR2673*. This example demonstrates that cleavage does not always take place between nucleotides 10 and 11 of the target transcript, thus providing a new reference for miRNA target prediction.

Because prediction methods cannot determine the authenticity of miRNA targets, experimental confirmation is required. To verify the accuracy of miRNA target prediction in our study, we compared the results of prediction analysis and degradome sequencing. Surprisingly, most of the predicted targets could not be verified by degradome sequencing. This was the case with several highly conserved longan miRNAs (*miR156*, *miR161*, *miR164*, *miR171*, *miR172*, *miR393*, *miR394*, and *miR408*) that were predicted to regulate homologs of known miRNA targets. Degradome sequencing failed to identify these targets, and also failed to identify other predicted targets for the above miRNAs. In addition, targets could not be predicted or verified for some highly conserved miRNAs using either prediction or degradome sequencing methods, although additional non-conserved targets were identified for them instead. This is consistent with the results of a previous study [36] and indicates that some conserved miRNAs had additional “novel” targets. Further analysis revealed that among 2063 predicted targets of 272 longan miRNAs, only 27 predicted targets of 18 miRNAs were verified by the degradome method. For example, 5 predicted targets of *miR2673*, including Unigene 49256 (expressed protein), Unigene 54104 (*GILI*), Unigene 60784 (similar to At1g73320), Unigene 67649 (similar to At4g28230), and Unigene 67737 (*MYB*), as well as several targets of *dlo-miR780*, *dlo-miR2099*, *dlo-miR5021*, and *dlo-miR5057*, were verified by degradome analysis (Table S4 in Supporting Information S1). There are several possible explanations for the observed discrepancies. For example, it is possible that miRNAs (e.g., *miR161*, *miR393*, and *miR394*) or targets were not detected because of their low expression levels during longan SE or because they were specific to tissues that had not been analyzed. Another possibility is that the degradome or transcriptome data was incomplete for target analysis. Alternatively, some miRNAs, such as *miR156* and *miR172*, inhibit expression of their target genes (*SBP* and *AP2*), primarily through translational arrest. Our results suggest that a combined prediction and degradome analysis approach is the most effective way to identify miRNA targets in non-model plants.

Validation of Longan *miR398* Targets Using Modified RLM-RACE

Recent studies have demonstrated that *miR398* plays a key role in plant SE [11,15,16]. Although *miR398* influences plant response

to oxidative stress by down-regulating Cu/Zn-superoxide dismutase (*CSD*) expression [37], it is not known whether this miRNA’s regulatory role in plant SE also involves *CSD* down-regulation. To verify the nature of predicted *miR398* targets and to investigate possible regulation of *CSD* by *miR398* during longan SE, we performed a psRNATarget analysis and a modified RLM-RACE experiment.

The psRNATarget analysis indicated that *dlo-miR398* can target *DICSD1a* and *DICSD1b*. This result is consistent with previous studies in *Arabidopsis* and rice [30,38,39], indicating that these miRNA binding sites are highly conserved among different plant species. A modified RLM-RACE experiment was then conducted, which verified *DICSD2a* as a target of *dlo-miR398* (Fig. 2). Tags derived from cleavage of *DICSD2a* by *dlo-miR398* were predominant, suggesting that *dlo-miR398* plays a major role in *DICSD2a* regulation during longan SE. On the other hand, the conserved target *DICSD1a* was not detected, implying that *DICSD1a* expression is not regulated by *dlo-miR398* during longan SE. These results indicate that *miR398* regulates longan SE primarily through mediation of the target *DICSD2a*.

Expression Patterns of Conserved and Novel miRNAs During Longan SE

Expression patterns of miRNAs were closely related to their functions. To better understand the role of miRNAs during longan SE, expression patterns of 24 miRNAs (20 conserved and 4 novel), *U6 snRNA*, and *5S rRNA* were analyzed using quantitative real-time PCR (qPCR). The qPCR analysis revealed that all selected miRNAs were expressed at varying levels in different embryogenic tissues, with some preferentially expressed in specific tissues. The miRNAs could be divided into roughly 3 groups (Fig. 3 and Fig. 4).

The first group, consisting of *U6 snRNA* and miRNAs, *dlo-miR159*, *dlo-miR4a*, *dlo-miR808*, and *dlo-miR5077*, displayed ubiquitous but varied expression levels during middle to late stages of longan SE. Different members of the *dlo-miR159* family had contrasting expression patterns, suggesting that each member of a family may have its own function. Expression of *miR159a.1* was high in heart- and torpedo-shaped embryos, but low in other tissues tested, especially in friable-embryogenic calli. *miR159f* was strongly expressed in heart-shaped and mature embryos. *miR159a.2* and *miR159c* had similar expression patterns during longan SE, being expressed weakly during early stages but abundantly in middle and late stages. *dlo-miR4* showed strong expression in torpedo-shaped and mature embryos, moderate expression in globular and heart-shaped embryos, and weak expression elsewhere. *dlo-miR5077* was highly expressed during late stages of longan SE. In addition, *dlo-miR808* and *U6 snRNA* also exhibited a similar expression pattern: unlike the other miRNAs, they were highly expressed during both early and late developmental stages, suggesting a regulatory role in morphogenesis during longan early and late SE.

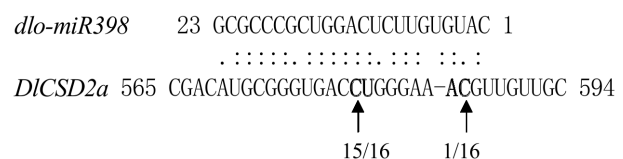


Figure 2. Cleavage site mapping of *miR398* target gene. mRNA sequence of *DICSD2a* is aligned with *miR398*. Nucleotides flanking the cleavage site are in bold. Numbers indicate the fraction of cloned PCR products terminating at different positions. doi:10.1371/journal.pone.0060337.g002

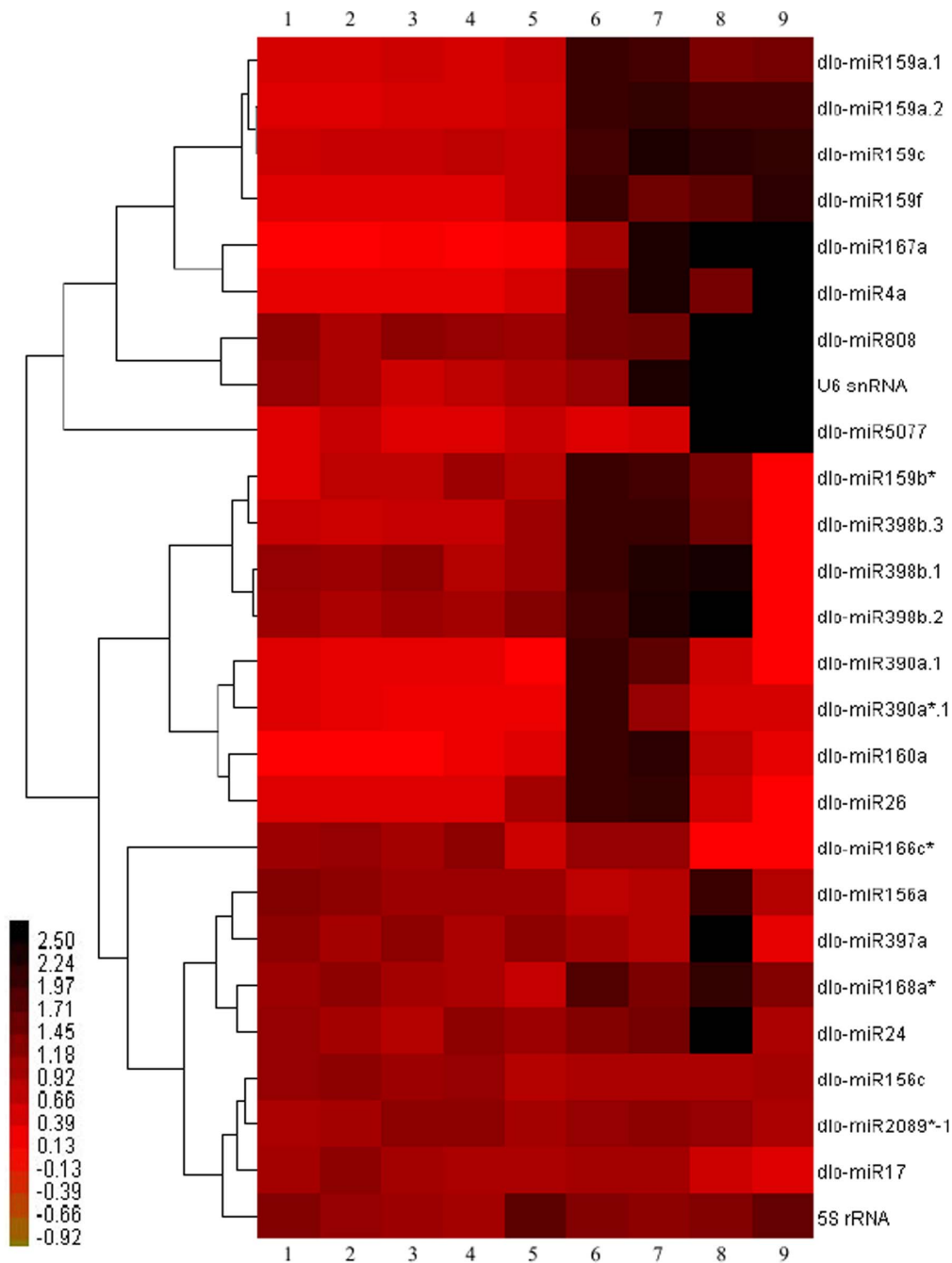


Figure 3. qPCR analysis of relative expressions of known and novel miRNAs, 5S rRNA, and U6 snRNA during longan SE. The bar represents the scale of relative expression levels of miRNAs, and colors indicate relative signal intensities of miRNAs. Each column represents a sample, and each row represents a single miRNA. Samples: 1. friable-embryogenic callus(EC); 2. embryogenic callus II(EC II); 3. incomplete compact pro-embryogenic cultures(ICpEC); 4. compact pro-embryogenic cultures(CpEC); 5. globular embryos(GE); 6. heart-shaped embryos(HE); 7. torpedo-shaped embryos(TE); 8. cotyledonary embryos(CE); 9. mature embryos(ME). doi:10.1371/journal.pone.0060337.g003

In the second category, *dlo-miR159b**, *dlo-miR160a*, *dlo-miR167a*, *dlo-miR390*, *dlo-miR398b*, and *dlo-miR26* exhibited tissue-specific expression, indicating that they may have stage-specific functions during longan SE. Among these, *dlo-miR159b** expression patterns were similar to those of the *miR398b* family (b.1, b.2, and b.3): they were all highly expressed in heart-shaped, torpedo-shaped, and cotyledonary embryos, but barely detectable in mature embryos. In an earlier study in oranges [16], *miR398* also exhibited high

expression in globular and cotyledonary embryos, suggesting that its function may be conserved across different species. *dlo-miR167a* was barely detectable during early longan SE stages; during late developmental stages, however, it was strongly expressed, suggesting a major role in SE during cotyledonary and mature embryonic stages. This is consistent with previous studies [16]. In addition, *dlo-miR390a.1* and *dlo-miR390a*.1*, differing by only 2 mismatches, showed low expression during early and late SE stages but high

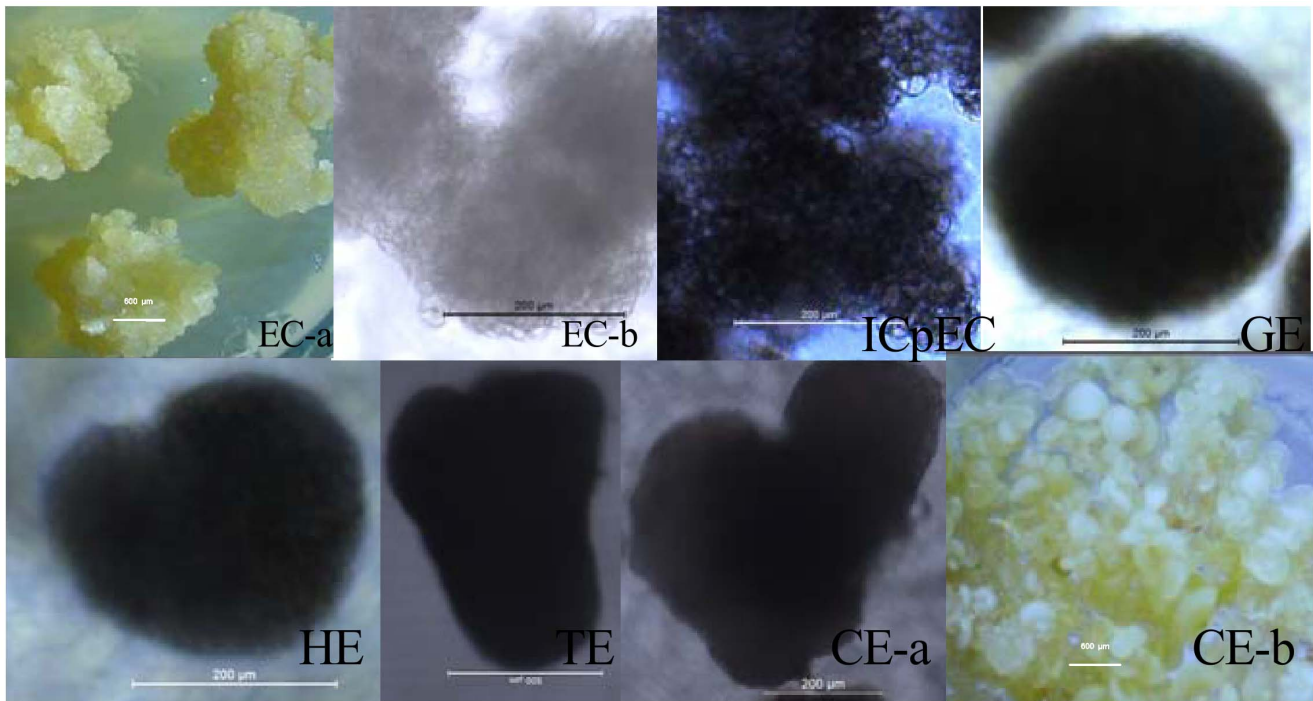


Figure 4. Morphology of embryogenic calli and embryos during the six sequential developmental stages of longan SE. Developmental stages are indicated at the left of each row. The bars in each phenotypic class are indicated at the middle of each image. The morphology of embryogenic cultures (EC-b, ICpEC, GE, HE, TE, CE-a) were observed using an inverted Leica DMIL LED microscope, except for EC-a and CE-b (bar = 200 μ m), the bars of others are 600 μ m; the images of EC-a and CE-b were both obtained under a Leica DFC295, bars, 600 μ m; EC, ICpEC, GE, HE, and TE were cultured on MS medium supplemented with 1 mg/L, 0.5 mg/L, 0.1 mg/L, 0.06 mg/L and 0.03 mg/L 2,4-D, respectively; and the CE was cultured on MS medium.

doi:10.1371/journal.pone.0060337.g004

expression during middle stages; *dlo-miR390a.1* was not detected in longan globular and mature embryos, but was expressed in orange globular embryos [16]. *dlo-miR160a* and *dlo-miR26* were highly expressed in the middle stages of SE, implying that their accumulation may be required for heart- and torpedo-shaped embryo formation and morphogenesis. On the other hand, *dlo-miR160a* was barely detectable during the period from the friable-embryogenic callus stage (stage 1) to the incomplete compact proembryogenic culture stage (stage 3), and *dlo-miR26* was not detected in mature embryos; this suggests that their targets of regulation were consistently promoted.

Members of the third group, comprising *dlo-miR156*, *dlo-miR166c**, *dlo-miR 168a**, *dlo-miR397a*, *dlo-miR 2089*-1*, *dlo-miR24*, *dlo-miR17*, and *5S rRNA*, were expressed at varied levels in different embryogenic tissues, suggesting their wide involvement in various developmental stages during longan SE. *dlo-miR156a* and *dlo-miR156c* were both highly expressed during early stages of longan SE. *dlo-miR156a* was strongly expressed in cotyledonary embryos, whereas *dlo-miR156c* expression was concentrated in embryogenic calli II (stage 2). *dlo-miR166c** showed high expression during friable-embryogenic callus through compact proembryogenic culture stages, moderate expression in globular embryos, and weak expression in cotyledonary and mature embryos; this indicates it may play a key role in early stages of longan SE. In contrast, *miR166* has been found to be abundantly expressed in orange cotyledonary embryos [16] and *Pinus taeda* zygotic embryos [18]. *dlo-miR 168a** expressed relatively high levels from heart-shaped embryos to mature embryos, especially in cotyledonary embryos, and high expression in cotyledonary embryos and moderate expression elsewhere was observed for

both *dlo-miR397a* and *dlo-miR24*. In addition to *dlo-miR17*, which was weakly expressed during late developmental stages, the remaining miRNAs *dlo-miR2089*-1* and *5S rRNA* showed ubiquitous expression during longan SE.

In summary, the above results indicate that different miRNA families and their individual members, as well as miRNAs*, have different regulatory roles during longan SE.

Discussion

In this study, 29 novel and 643 conserved miRNAs in more than 169 miRNA families were identified during longan SE by Solexa sequencing. This research represents the first large-scale cloning and characterization of plant SE-related miRNAs. Although many miRNAs remain to be identified, our results reveal that longan has a larger and more diverse miRNA (snRNA) population than other plants [11,15,16] when conserved miRNAs are compared. Bioinformatics prediction, degradome analysis, and modified RLM-RACE demonstrated that putative miRNA targets were involved in a broad variety of biological processes, including plant metabolism, signal transduction, and stimulus response. The stage- and tissue-specific expression patterns observed for 24 conserved and novel miRNAs suggest their probable roles in longan SE, and imply that different miRNA families are responsible for morphogenesis at different developmental stages.

Putative Functions of miRNAs during Longan SE Based on Analysis of Expression Patterns

SE is a multi-step regeneration process which begins with proembryogenic mass formation followed by somatic embryo

formation, maturation, desiccation, and plant regeneration. This complicated process involves a large number of genes with complex expression patterns. By identifying miRNAs and their targets and studying their expression patterns, we can propose possible roles for miRNAs during longan SE.

The largest miRNA family in plants—*miR156*—plays critical regulatory roles during juvenile–adult transitions [40] and flowering [41] by targeting *SPL* genes. In *Arabidopsis*, beginning as early as the eight-cell stage, *miR156*-mediated repression of zygotic *SPL10* and *SPL11* prevents premature accumulation of transcripts from genes normally induced during the embryonic maturation phase [14]. In the study, the highest concentrations of *dlo-miR156a* were observed in cotyledonary embryos, which is consistent with *miR156* patterns in larch [15]. In contrast, *dlo-miR156c* expression was strongest in embryogenic calli II (stage 2), similar to results in citrus, in which *miR156* levels were at a maximum during the E2 (EC induced for 2 weeks) stage [16]. These results suggest that *miR156* members play regulatory roles during early stages of longan SE or cotyledonary embryo development by targeting *SPL*, and further imply that functional diversification can take place through the existence of multicopy miRNA families.

In *Arabidopsis*, ABA induction of *miR159* controls transcript levels of 2 *MYB* factors during seed germination [42]. In our study, *dlo-miR159* family transcripts accumulated during middle and late stages of longan SE. This result is similar to findings with respect to *miR159* at the cotyledonary embryo stage of larch [15]. It is not consistent, however, with the pattern observed in citrus, in which *miR159* was expressed at significantly higher levels in globular-shaped embryos [16]. These contrasting results suggest that the role of *miR159* differs between species, and imply that *miR159* may function in middle and late stages of longan SE by targeting Unigene 66155 (*MYB33*).

*dlo-miR159b**, *dlo-miR160a*, *dlo-miR167a*, *dlo-miR390*, *dlo-miR398b*, and *dlo-miR26* may have stage-specific functions, because they exhibit tissue-specific expression during longan SE. *dlo-miR159b**, the complementary strand of mature functional miRNA, exhibited an expression pattern similar to that of other *miR159* family members; unlike them, it was not expressed in mature cotyledonary embryos, demonstrating that it regulates different targets during longan SE. *miR160*, which regulates the expression of *ARF16* and *ARF17* during embryonic development in *Arabidopsis* [43], was barely detectable during early stages of longan SE. It was instead most highly expressed during heart- and torpedo-shaped embryonic stages, suggesting that *dlo-miR160a* is involved in longan heart- and torpedo-shaped embryo formation and morphogenesis. This also contrasts with larch *miR160*, which may play a regulatory role during cotyledonary embryo development. *miR167* regulates plant fertility by targeting *ARF6* and *ARF8* in response to auxin signal [44–46]. In our study, the strongly-expressed *miR167a* appears to play a major role during cotyledonary and mature embryonic stages by regulating *ARFs* (Unigenes 65933/62178, *ARF3/8*), which is consistent with previous studies in larch and oranges [15,16]. In rice, *miR167* levels gradually decrease when cells cultured in the presence of auxin are transferred to an auxin-free medium [47]. Surprisingly in our study, *dlo-miR167a* was undetectable in a medium containing 2,4-D, perhaps because of the endogenous hormone levels present in the longan embryogenic cultures. *miR167* can be regulated by ABA and GA_3 [48]. A previous study found that during longan SE, the ratio of endogenous ABA/ GA_3 gradually increases [49]; in our study, the expression trend of *dlo-miR167a* generally mirrored that of ABA/ GA_3 , indicating that *dlo-miR167a* accumulates in response to endogenous auxin, and that reduced

expression of its targets is required for cotyledonary and mature embryo morphogenesis in longan. Like *miR160*, *dlo-miR390 a.1* and *-a*.1* also accumulated during heart- and torpedo- shaped embryonic stages. Previous studies have shown that *miR390* directs the formation of ta-siRNAs, which target *ARF2*, *ARF3*, and *ARF4* genes to modulate developmental timing and patterning in *Arabidopsis* [50], and that this miRNA is most highly expressed in citrus in globular-shaped embryos [16] and in larch in cotyledonary embryos [15]. These differing results among different species may be a consequence of different *miR390* members being examined, and suggest that a strict division of labor may exist in this miRNA family.

During plant embryonic development, cells not only increase in number, but also undergo differentiation. By inducing autonomous cell division, oxidative stress can modulate plant SE [51,52]. *miR398* is the first miRNA to be directly linked to regulation of plant oxidative stress response via down-regulation of *CSD* expression [37]. Recent studies have shown that *miR398* is required for cotyledon-shaped embryo morphogenesis in citrus [16] and is involved in modulation of proembryogenic mass propagation and transition to single embryos in larch [15]. Surprisingly, *dlo-miR398a* was barely detected during longan SE in our study (data not shown). Moreover, in a previous study, *pre-miR398a* was only weakly expressed during early longan SE stages and was not detected at all during middle and late stages [53]. These results indicate that *dlo-miR398a* is not a key developmental factor in longan SE. In addition, the low expression of *dlo-miR398a* led to increased accumulation of its target (*DICSD1a*), such that mRNA cleavage barely took place; this may explain why cleaved *DICSD1a* mRNAs were not detected using modified RLM-RACE in our study. In contrast to *dlo-miR398a*, the *dlo-miR398b* family was highly expressed in heart-shaped, torpedo-shaped, and cotyledonary embryos, suggesting its involvement in these developmental stages. *miR398* targets *CSD* and is down-regulated under stress conditions [39]. In our study, the addition of 5% sucrose to the mature embryo induction medium produced a stress-like osmotic pressure. Low expression levels of *dlo-miR398b* during cotyledonary embryo maturation permitted accumulation of *DICSD2a* as a response to this stress, promoting maturation of longan cotyledonary embryos, consistent with a previous report [16].

High expression in cotyledonary embryos and moderate expression elsewhere was observed for both *dlo-miR397a* and *dlo-miR24*. *miR397* is known to target laccases, which are associated with cell wall lignification and thickening during secondary cell growth [54]. In citrus, *miR397* is specifically expressed at high levels in globular embryos; in larch, *miR397* levels are highest in early cotyledonary embryos—consistent with our results—with low levels observed for its target mRNA. We consequently propose that *dlo-miR397a* helps regulate cell wall thickness during SE through cleavage of laccase mRNA.

miR166 targets HD-ZIP III transcription factors [55] and is required for cotyledon-shaped embryo morphogenesis [15,16]. In our study, similar targets, such as Unigenes 12408 and 14044 (class III HD-Zip proteins 4 and 5), were also predicted for *dlo-miR166*, demonstrating the evolutionarily conserved nature of these miRNA targets in different plants. In another study, *miR166* was induced by gibberellic acid (GA_3) [56]. Expression levels of endogenous GA_3 showed a gradual downward trend during longan SE [49], which was mirrored in our study by *dlo-miR166c** expression levels. This suggests that changes in *miR166c** levels are caused by alterations in endogenous GA_3 concentrations. GA_3 inhibits the early stage of embryogenic cell differentiation/development leading to globular embryos [57]; *dlo-miR166c** in

our study exhibited a decreasing level of expression during this transition, suggesting a role for *dlo-miR166** during early stages of longan SE. *dlo-miR168a** expressed relatively high levels from stages heart-shaped embryos to mature embryos, especially in cotyledonary embryos, which is consistent with *miR168* from larch [15]. *miR168* is associated with repression of *AGO1* accumulation, and *AGO1* is the most important AGO protein in the miRNA pathway, which has specific expression during carrot SE [19]. Therefore, we speculated the accumulation of *dlo-miR168a** during cotyledonary embryos development repressed the expression of *AGO1*, thus leading to the development of longan embryo.

In addition, several miRNAs, such as *dlo-miR808*, *-2089** and *-5077* were also detected in longan SE, which have rarely been studied in plant development. Among them, *dlo-miR808* and *-5077* both showed highly expression embryos at late developmental stages of longan somatic embryo. the former reaches an abundance peak at cotyledonary embryos; and the later, which is similar to *osa-miR5077*, reaches peak at mature embryos, suggesting the accumulation of the two miRNAs is required for the development of somatic embryo at late stages. Besides, *miR2089**, which considered to be legume-specific [58], showed ubiquitous expression levels and exerted regulatory function during longan SE. Further, the expression analysis of the 4 novel miRNAs uncovered in this study (*dlo-miR4a*, *dlo-miR17*, *dlo-miR26*, and *dlo-miR24*), showed varied levels in different embryogenic tissues during longan SE, has laid a foundation for future functional studies of longan embryogenesis. Such a study is currently underway in our laboratory. The investigation reported here represents a small but significant step towards the elucidation of functions of miRNAs during longan embryogenesis.

Materials and Methods

Plant Materials and RNA Isolation

Synchronized embryogenic cultures at different developmental stages were obtained following previously published methods for longan [1,2,58–60]. The synchronized cultures, consisting of friable-embryogenic calli, embryogenic calli II, incomplete compact pro-embryogenic cultures, compact pro-embryogenic cultures, globular embryos, heart-shaped embryos, torpedo-shaped embryos, cotyledonary embryos, and mature embryos, were collected and stored at -80°C for later use. Total RNAs were extracted from the above-described cultures using Trizol reagent (Invitrogen, Carlsbad, CA, USA). Only RNA samples with *A260/A280* ratios between 1.9 and 2.1 and *A260/A230* ratios higher than 2.0 were used for further analyses.

Longan sRNA cDNA Library Construction, Solexa Sequencing, and Data Analysis

To identify potential conserved and novel miRNAs from longan SE, a pooled sRNA library was generated from the synchronized embryogenic cultures (friable-embryogenic calli, globular embryos, heart-shaped embryos, torpedo-shaped embryos, and cotyledonary embryos) and sequenced on a Solexa system (Illumina). First, small (16–30 nt) RNA fragments were isolated from a 15% PAGE gel and purified. The small RNAs were then sequentially ligated to a 5' RNA adapter (5'-GUUCAGAGUUCUACAGUCCGAC-GAUC-3') and a 3' RNA adapter (5'-pUCG UAUGCCGU-CUUCUGCUUGidT-3'; p, phosphate; idT, inverted deoxythymidine) using T4 RNA ligase. The resulting samples were reverse-transcribed to cDNA with an RT primer (5'-CAAGCAGAA-GACGGCATAACGA-3') using Superscript II reverse transcriptase (Invitrogen) and PCR-amplified. Finally, Solexa sequencing was

used to sequence the sRNAs in the pooled longan samples (Beijing Genomics Institute, China).

Bioinformatics analyses were conducted on the resulting Solexa sequencing data. The 35-nt sequence tags were first trimmed of adaptors, regions of low complexity, and low-quality sequences, and the length distribution of clean tags was summarized. The remaining sRNA sequences (clean reads) were mapped to both a longan embryogenic callus transcriptome (SRA050205) and a *P. trichocarpa* genome. To classify degradation fragments of noncoding RNA, the clean tags were then compared against non-coding RNAs in Rfam and NCBI GenBank databases. Any sRNAs having exact matches to these sequences were excluded from further analysis. To identify conserved miRNAs in longan, the unique sRNA sequences were searched using miRBase 17.0. Novel miRNAs were predicted from unannotated sRNAs using Mireap, and their secondary structures were predicted using the computational software package Mfold 3.1. At the same time, first position base bias among sRNA candidates of certain lengths was summarized to determine prediction accuracy. Finally, to understand the biological functions of miRNAs, potential longan novel and conserved miRNA targets were predicted as described in [41,61].

Degradome cDNA Library Construction and Sequencing

To experimentally verify both conserved and novel miRNA targets in longan, a degradome cDNA library from the above pooled total RNA samples was constructed using the following steps: ligation of polyA-enriched RNA samples to a custom RNA adaptor containing an *MmeI* site, followed by reverse transcription (RT), second-strand synthesis, *MmeI* digestion, ligation of a 3' double-stranded DNA adaptor, gel purification, and PCR amplification. The amplified degradome tag library was sequenced on a Solexa/Illumina genome analyzer. After sequencing, 20- and 21-nt reads from the degradome library were analyzed using PairFinder to identify sRNA targets.

psRNATarget and Modified 5' RLM-RACE

To verify the nature of predicted *miR398* targets and to investigate possible regulation of *CSD* by *miR398* during longan SE, a psRNATarget [62] analysis and a modified 5' RLM-RACE experiment were performed. For the psRNATarget analysis, sequences of the newly-identified longan *miR398* were matched against nucleic acid sequences of longan *SOD* gene families (*Cu/Zn-SOD*, *Fe-SOD*, and *Mn-SOD*). All predicted targets were scored using criteria described in [62]. Any sequences with total scores less than 5.0 were considered to be miRNA targets. To identify cleavage sites of target transcripts, a modified RLM-RACE experiment was performed using a GeneRacer kit (Invitrogen). Briefly, total RNAs from equal mixtures of friable-embryogenic calli, globular embryos, heart-shaped embryos, and torpedo-shaped embryos were ligated to a 5' RACE RNA adapter. The resulting samples were reverse-transcribed to cDNA using a GeneRacer Oligo dT primer and Superscript III reverse transcriptase, and PCR-amplified using GeneRacer 5' primers and gene-specific primers (Table S5 in Supporting Information S1). The resulting PCR products were gel-purified, cloned, and sequenced.

Real-time Quantitative PCR of Longan Conserved and Novel miRNAs

qPCR was used to validate results obtained from high-throughput sequencing of longan miRNAs. RNA samples from the 9 embryogenic cultures described above were reverse-transcribed using an NCode VILO miRNA cDNA Synthesis kit. Expression profiles of 24 miRNAs were examined using an NCode

Express SYBR GreenER miRNA qPCR kit (Invitrogen). All reactions were performed in triplicate in a LightCycler 480 qPCR instrument (Roche Applied Science, Switzerland), with a dissociation curve used to control for primer dimers in the reactions. Mature miRNA abundance was calculated relative to expression of reference genes *dlo-miR156c*, *dlo-miR2089*-1*, and *5S rRNA*. miRNA names and primer sequences are provided in Table S6 in Supporting Information S1.

Supporting Information

Supporting Information S1 File containing the following supporting information tables. Table S1. Conserved and non-

conserved longan miRNAs. Table S2. Longan-known miRNA targets by prediction method. Table S3. Identification targets for longan miRNAs by Degradome sequencing. Table S4. The consensus targets for longan miRNAs obtained from prediction and degradome analysis. Table S5. Primer sequences. Table S6. The miRNA names, primer sequences.

(ZIP)

Author Contributions

Conceived and designed the experiments: ZL YL. Performed the experiments: ZL YL. Analyzed the data: ZL YL. Contributed reagents/materials/analysis tools: ZL YL. Wrote the paper: ZL YL.

References

- Lai Z, Chen C, Chen Z (2001) Progress in biotechnology research in longan. *Acta Horticulturae* 58: 137–141.
- Lai Z, Chen C, Zeng L, Chen Z (2000) Somatic embryogenesis in longan (*Dimocarpus longan* Lour.). *Forestry Sciences*, Dordrecht: Kluwer Academic Publishers 67: 415–432.
- Lai ZX, Chen ZX, He BZ (1998) Somatic embryogenesis in longan. Larkin PJ *Agricultural Biotechnology: Laboratory, Field and Market-Proceedings of the 4th Asia-Pacific Conference on Agricultural Biotechnology* Canberra: UTC Publishing 369–371.
- Lai ZX, He Y, Chen YT, Cai YQ, Lai CC, et al. (2010) Molecular biology and proteomics during somatic embryogenesis in *Dimocarpus longan* Lour. *Acta Hort (ISHS)* 863.
- Lin YL, Lai ZX (2010) Reference gene selection for qPCR analysis during somatic embryogenesis in longan tree. *Plant Science* 178: 359–365.
- Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 297: 2053–2056.
- Chuck G, Candela H, Hake S (2009) Big impacts by small RNAs in plant development. *Curr Opin Plant Biol* 12: 81–86.
- Chen X (2005) MicroRNA biogenesis and function in plants. *FEBS Lett* 579: 5923–5931.
- Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. *Science* 301: 336–338.
- Bartel D (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281–297.
- Luo YC, Zhou H, Li Y, Chen JY, Yang JH, et al. (2006) Rice embryogenic calli express a unique set of microRNAs, suggesting regulatory roles of microRNAs in plant post-embryonic development. *FEBS Lett* 580: 5111–5116.
- Mica E, Piccolo V, Delledonne M, Ferrarini A, Pezzotti M, et al. (2009) High throughput approaches reveal splicing of primary microRNA transcripts and tissue specific expression of mature microRNAs in *Vitis vinifera*. *BMC Genomics* 10: 558.
- Willmann MR, Mehalick AJ, Packer RL, Jenik PD (2011) MicroRNAs regulate the timing of embryo maturation in *Arabidopsis*. *Plant Physiol* 155: 1871–1884.
- Nordine MD, Bartel DP (2010) MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. *Genes Dev* 24: 2678–2692.
- Zhang J, Zhang S, Han S, Wu T, Li X, et al. (2012) Genome-wide identification of microRNAs in larch and stage-specific modulation of 11 conserved microRNAs and their targets during somatic embryogenesis. *Planta* 236: 647–657.
- Wu XM, Liu MY, Ge XX, Xu Q, Guo WW (2011) Stage and tissue-specific modulation of ten conserved miRNAs and their targets during somatic embryogenesis of Valencia sweet orange. *Planta* 233: 495–505.
- Zhang S, Zhou J, Han S, Yang W, Li W, et al. (2010) Four abiotic stress-induced miRNA families differentially regulated in the embryogenic and non-embryogenic callus tissues of *Larix leptolepis*. *Biochem Biophys Res Commun* 398: 355–360.
- Oh TJ, Wartell RM, Cairney J, Pullman GS (2008) Evidence for stage-specific modulation of specific microRNAs (miRNAs) and miRNA processing components in zygotic embryo and female gametophyte of loblolly pine (*Pinus taeda*). *New Phytol* 179: 67–80.
- Takahata K (2008) Isolation of carrot Argonaute1 from subtractive somatic embryogenesis cDNA library. *Biosci Biotechnol Biochem* 72: 900–904.
- Liang C, Zhang X, Zou J, Xu D, Su F, et al. (2010) Identification of miRNA from *Porphyra yezoensis* by high-throughput sequencing and bioinformatics analysis. *PLoS One* 5: e10698.
- Zhao CZ, Xia H, Frazier TP, Yao YY, Bi YP, et al. (2010) Deep sequencing identifies novel and conserved microRNAs in peanuts (*Arachis hypogaea* L.). *BMC Plant Biol* 10: 3.
- Qiu D, Pan X, Wilson IW, Li F, Liu M, et al. (2009) High throughput sequencing technology reveals that the taxoid elicitor methyl jasmonate regulates microRNA expression in Chinese yew (*Taxus chinensis*) *Gene* 436: 37–44.
- Song C, Wang C, Zhang C, Korir NK, Yu H, et al. (2010) Deep sequencing discovery of novel and conserved microRNAs in trifoliate orange (*Citrus trifoliata*). *BMC Genomics* 11: 431.
- Joshi T, Yan Z, Libault M, Jeong DH, Park S, et al. (2010) Prediction of novel miRNAs and associated target genes in *Glycine max*. *BMC Bioinformatics* 11 Suppl 1: S14.
- Moxon S, Jing R, Szittyá G, Schwach F, Rusholme Pilcher RL, et al. (2008) Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Res* 18: 1602–1609.
- Yao Y, Guo G, Ni Z, Sunkar R, Du J, et al. (2007) Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.). *Genome Biol* 8: R96.
- Barakat A, Wall PK, Diloreto S, Depamphilis CW, Carlson JE (2007) Conservation and divergence of microRNAs in *Populus*. *BMC Genomics* 8: 481.
- Wu L, Zhou H, Zhang Q, Zhang J, Ni F, et al. (2010) MicroRNA Mediates DNA Methylation. *Molecular Cell*, 38: 465–475.
- Jones-Rhoades MW, Bartel DP, Bartel B (2006) MicroRNAs and their regulatory roles in plants. *Annual Review of Plant Biology* 57: 19–53.
- Jones-Rhoades MW, Bartel DP (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol Cell* 14: 787–799.
- Dezulian T, Palatnik JF, Huson D, Weigel D (2005) Conservation and divergence of microRNA families in plants. *Genome Biology* 2005, 6: P13.
- Puzey JR, Karger A, Axtell M, Kramer EM (2012) Deep annotation of *Populus trichocarpa* microRNAs from diverse tissue sets. *PLoS One* 7: e33034.
- Pantaleo V, Szittyá G, Moxon S, Miozzi L, Moulton V, et al. (2010) Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. *Plant J* 62: 960–976.
- Li YF, Zheng Y, Addo-Quaye C, Zhang L, Saini A, et al. (2010) Transcriptome-wide identification of microRNA targets in rice. *Plant J* 62: 742–759.
- Zhou M, Gu L, Li P, Song X, Wei L, et al. (2010) Degradome sequencing reveals endogenous small RNA targets in rice (*Oryza sativa* L. ssp. indica). *Frontiers of Biology* 5: 67–90.
- Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ (2008) Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis* degradome. *Curr Biol* 18: 758–762.
- Sunkar R, Li YF, Jagadeeswaran G (2012) Functions of microRNAs in plant stress responses. *Trends Plant Sci* 17: 196–203.
- Bonnet E, Wuys J, Rouze P, Van de Peer Y (2004) Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes. *Proc Natl Acad Sci U S A* 101: 11511–11516.
- Sunkar R, Kapoor A, Zhu JK (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in *Arabidopsis* is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell* 18: 2051–2065.
- Cardon GH, Hohmann S, Nettesheim K, Saedler H, Huijser P (1997) Functional analysis of the *Arabidopsis thaliana* SBP-box gene SPL3: a novel gene involved in the floral transition. *Plant J* 12: 367–377.
- Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, et al. (2005) Specific effects of microRNAs on the plant transcriptome. *Dev Cell* 8: 517–527.
- Reyes JL, Chua NH (2007) ABA induction of miR159 controls transcript levels of two MYB factors during *Arabidopsis* seed germination. *Plant J* 49: 592–606.
- Liu X, Huang J, Wang Y, Khanna K, Xie Z, et al. (2010) The role of floral organs in carpels, an *Arabidopsis* loss-of-function mutation in MicroRNA160a, in organogenesis and the mechanism regulating its expression. *Plant J* 62: 416–428.
- Nagpal P, Ellis CM, Weber H, Ploense SE, Barkawi LS, et al. (2005) Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* 132: 4107–4118.
- Tabata R, Ikezaki M, Fujibe T, Aida M, Tian CE, et al. (2010) Arabidopsis auxin response factor6 and 8 regulate jasmonic acid biosynthesis and floral organ development via repression of class 1 KNOX genes. *Plant Cell Physiol* 51: 164–175.
- Wu MF, Tian Q, Reed JW (2006) Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development* 133: 4211–4218.

47. Yang JH, Han SJ, Yoon EK, Lee WS (2006) Evidence of an auxin signal pathway, microRNA167-ARF8-GH3, and its response to exogenous auxin in cultured rice cells. *Nucleic Acids Res* 34: 1892–1899.
48. Zhang BH, Pan XP, Wang QL, Cobb GP, Anderson TA (2005) Identification and characterization of new plant microRNAs using EST analysis. *Cell Res* 15: 336–360.
49. Lai ZX, Chen CL (2002) Changes of endogenous phytohormones in the process of somatic embryogenesis in longan (*Dimocarpus longan* Lour.). *Chinese Journal of Tropical Crops* 23: 41–47.
50. Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, et al. (2006) Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in *Arabidopsis*. *Curr Biol* 16: 939–944.
51. Kairong C, Ji L, Gengmei X, Jianlong L, Lihong W, et al. (2002) Effect of hydrogen peroxide on synthesis of proteins during somatic embryogenesis in *Lycium barbarum* Plant Cell, Tissue and Organ Culture 68: 187–193.
52. Pasternak TP, Prinsen E, Ayaydin F, Miskolczi P, Potters G, et al. (2002) The Role of auxin, pH, and stress in the activation of embryogenic cell division in leaf protoplast-derived cells of alfalfa. *Plant Physiol* 129: 1807–1819.
53. Lin YL, Lai ZX (2011) Cloning of miR398a precursor gene from embryogenic callus and its expression analysis during somatic embryogenesis in longan. *Chinese journal of Tropical Crops* 4: 632–639.
54. Constabel CP, Yip L, Patton JJ, Christopher ME (2000) Polyphenol oxidase from hybrid poplar. Cloning and expression in response to wounding and herbivory. *Plant Physiol* 124: 285–295.
55. Williams L, Grigg SP, Xie M, Christensen S, Fletcher JC (2005) Regulation of Arabidopsis shoot apical meristem and lateral organ formation by microRNA miR166 g and its AtHD-ZIP target genes. *Development* 132: 3657–3668.
56. Liu Q, Zhang YC, Wang CY, Luo YC, Huang QJ, et al. (2009) Expression analysis of phytohormone-regulated microRNAs in rice, implying their regulation roles in plant hormone signaling. *FEBS Lett* 583: 723–728.
57. Tokuji Y, Kuriyama K (2003) Involvement of gibberellin and cytokinin in the formation of embryogenic cell clumps in carrot (*Daucus carota*). *J Plant Physiol* 160: 133–141.
58. Lai ZX, Chen ZG (1997) Somatic embryogenesis of high frequency from longan embryogenic calli. *Journal of Fujian Agricultural University* 26: 271–276.
59. Chen CL, Lai ZX (2002) Synchronization regulation of embryogenesis of embryogenic calli and their histological observations in longan. *Journal of Fujian Agriculture and Forestry University (Natural Science Edition)*, 31: 192–194.
60. Fang ZZ, Lai ZX, Chen CL (2009) Preliminary synchronization regulation at the middle developmental stage during longan somatic embryogenesis. *Chinese Agricultural Science Bulletin* 25: 152–155.
61. Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121: 207–221.
62. Zhang Y (2005) miRU: an automated plant miRNA target prediction server. *Nucleic Acids Res* 33: W701–704.