Identification of miRNA encoded by *Jatropha curcas* from EST and GSS

Nutan Prakash Vishwakarma1,* and Vasant J. Jadeja2

1Department of Biotechnology; Shree M. & N. Virani Science College; Rajkot, Gujarat, India; 2Department of Microbiology; Shree M. & N. Virani Science College; Rajkot, Gujarat, India.

Keywords: microRNA, comparative genomics, expressed sequence tags, genome survey sequence, *Jatropha curcas*

Abbreviations: ncRNA, Non coding RNA; miRNA, microRNA; Pre-miRNA, Precursor microRNA; miRNA*, passenger strand on precursor miRNA; MFE, minimal free energy; AMFE, adjusted MFE; MFEi, minimal free energy index; EST, expressed sequence tag; GSS, genome survey sequence

miRNAs are endogenous approx 22 nucleotide RNA which mediates transcriptional or Post-transcriptional gene regulation and play a critical role in diverse aspects of plant development. miRNA identification in wet lab have various constraints, it is time consuming and expensive. It also faces the limitation of identifying miRNAs expressed at specific time and/or at special conditions. Due to the nature of strong conservation of miRNA in plant species, the use of comparative genomics approach for expressed sequence tags (ESTs), Genome Survey Sequence (GSS) and structural feature criteria filter has paved the way toward the identification of conserved miRNAs from the plant species whose genomes are not yet available in public domain. To identify the novel miRNA from *Jatropha curcas*, a total of 46862 EST sequences and 1569 GSS were searched for homology to previously known viridiplantae 2502 mature miRNA. After predicting the RNA secondary structure, 24 new potential miRNA were identified in *J. curcas*. Using the newly identified miRNA sequences, a total of 78 potential target genes were identified for 3 miRNA families. Most of the miRNA targeted genes were predicted to encode transcription factors that regulate cell growth and development, signaling, and metabolism. These findings considerably broaden the scope of understanding the functions of miRNA in *J. curcas*.

Introduction

The discovery of miRNA played a very important role in the field of biological research.1,2 The first Plant miRNAs were identified after a long time in early 2002 from *Arabidopsis thaliana*.3 Since then thousands of plant miRNA have been submitted in miR-Base Sequence database (http://www.mirbase.org).4 miRNAs are endogenous approx 21–24 nucleotide RNA which mediates transcriptional or post-transcriptional gene regulation and play a critical role in diverse aspects of plant development including auxin signaling, meristem boundary formation, organ separation, leaf development and polarity, lateral root formation, transition from juvenile to adult vegetative phase and from vegetative to flowering phase, floral organ identity and reproduction, as well as adaptation to biotic and abiotic stresses, including nutrient deprivation.5,6

Plant miRNAs are primarily found in genomic regions not associated with protein-coding genes and most of the plant miRNAs are produced from their own transcriptional units. miRNA genes are transcribed by RNA polymerase II (pol II). The primary miRNA transcripts (pri-miRNAs) contain cap structures as well as poly(A) tails.6 Mature miRNAs are produced from longer RNA hairpin precursors by the endoribonuclease III-like enzyme, dicer like-1 (DCL1). The processed and methylated miRNA/miRNA* duplex is exported to the cytosol via HASTY5, a plant ortholog of exportin.7 miRNAs that are incorporated into an argonaute containing RNA-induced silencing complex can affect the target gene expression. miRNA mediated regulations rely on specific miRNA target mRNA interactions that result in degradation of the target transcript and/or attenuation of translation.8,9

Although hundreds of miRNA have been discovered in recent years in plants many other miRNA gene functions remain to be elucidated. For the discovery of novel microRNA in different species four different methods are used.10 These are (1) genetic screening; (2) direct cloning through constructing a small RNA library; (3) traditional computational approach based on whole genome sequences; and (4) EST analysis.

Genetics screening technology is the first method for discovery of microRNAs.2 This technology requires the isolation of total RNA and sequencing each RNA and therefore it is time-consuming, complicated and expensive, and hence this method is not much practiced.10

*Correspondence to: Nutan Prakash Vishwakarma; Email: nutanp@vsc.edu.in
Submitted: 12/05/12; Accepted: 12/06/12
http://dx.doi.org/10.4161/psb.23152
Citation: Vishwakarma N, Jadeja V Identification of miRNA encoded by *Jatropha curcas* from EST and GSS. Plant Signal Behav 2013; 8:e23152; PMID: 23299511; http://dx.doi.org/10.4161/psb.23152.

*Correspondence to: Nutan Prakash Vishwakarma; Email: nutanp@vsc.edu.in
Submitted: 12/05/12; Accepted: 12/06/12
http://dx.doi.org/10.4161/psb.23152
Citation: Vishwakarma N, Jadeja V Identification of miRNA encoded by *Jatropha curcas* from EST and GSS. Plant Signal Behav 2013; 8:e23152; PMID: 23299511; http://dx.doi.org/10.4161/psb.23152.
In Direct cloning technology after isolation of small RNAs, the RNAs are separated by gel electrophoresis. Usually, a 5’ and/or 3’ adaptor is added to the small RNA for easily operating the small RNAs. Finally, the small RNAs are sequenced and the miRNAs are identified from the isolated small RNAs. This method has been widely employed to identify miRNAs from different species in plants and animals. The major limitation of this method is that it is time-consuming and expensive. This method also fails to identify miRNAs expressed at specific time and/or at special conditions.

In the case of availability of whole genome sequence of organism large number of computational approaches has been designed to search miRNA. The unavailability of whole genome sequences played a great limitation in this strategy of the miRNA identification.

EST analysis is a powerful tool to identify miRNAs conserved among various plant species whose whole genome sequences are not available, and to study the conservation and evolution of miRNAs among different species. EST analysis strategy have been proven to be successful for the discovery of new miRNAs from various plant species including Cotton, Tea, Oil palm, Radish, Potato, Citrus, rapeseed, Lettuce, Soybean, Wheat, Tobacco, Groundsel, and Cowpea. Thus, the computational or bioinformatics based approach is very useful for predicting novel miRNAs, which usually cannot be detected by the direct cloning.

Just like EST analysis, one more nucleotide database, GSS is also used in miRNA identification field and by this strategy, many miRNAs in several plant species are identified. Due to the excessive use of non-renewable hydrocarbons worldwide as energy source resulted it in fast depletion of reserves and threat of global warming, which the researcher made to search for renewable sources. Plant based biodiesel derived from the seed oil of \textit{J. curcas} is now emerging as an available alternative to the conventional fossil fuel. \textit{J. curcas} is a Latin American origin hardy perennial shrub, which is widespread throughout the tropical regions of the world. Jatropha is a large genus comprising more than 175 species among except for \textit{J. curcas} and \textit{J. glandulifera}, which are oil-yielding species, most of which is ornamental. Apart from being a potential biofuel crop \textit{J. curcas} is multipurpose species with many attributes and considerable potential.

Thousand of microRNAs have been discovered in recent years but there have been only 46 novel miRNAs that have been reported for the Jatropha by the cloning approach. Since during the cloning approach, there is a chance of degradation of miRNAs and skipping of loosely expressed miRNAs. Therefore, in the present study, we have used all known plant miRNAs (publicly available in miRBase) from viridiplantae to search the conserved \textit{J. curcas} miRNA homologs in publicly available EST and GSS databases.

### Results

A total of 20 four potential miRNAs were detected with predicted stem-loop precursor structure from the publically available EST and GSS database.

### Identification of potential \textit{J. curcas} miRNAs

| miRNA family | Predicted targets |
|--------------|------------------|
| miR166       | 46               |
| miR167       | 32               |
| miR1096      | 6                |
| miR5368      | 5                |
| miR5021      | 1                |

These miRNA gene candidates are predicted to encode Jatropha members of the jcu-miR166, jcu-miR167, jcu-miR1096, jcu-miR5368, jcu-miR5021 family. The identified potential miRNAs have both higher negative minimal fold energies (MFEs) and MFEIs (>0.87–4.58867) (Table 1). We have chosen only those miRNA candidates which shows the value of MFEI greater than 0.85, which is a commonly used value to distinguish miRNAs from other non-coding and coding RNAs.

Total 24 predicted miRNAs belongs to 5 different miRNA families in Jatropha as shown in Figure 2, miRNA family miR166, miR167, miR1096, miR5368 and miR5021 have 9, 9, 1, 2 and 3 members respectively. The different size of the identified miRNAs within different families suggests that they may offer unique functions for regulation of miRNA biogenesis or gene expression. Predicted miRNAs also depicts the diversity in the location of mature miRNA sequences. The sequences of miR166a/b/c/d/e/f/g/h/i and miR5021a/b/c were located at the 5’ end of the miRNA precursors, while the others were found at the 3’ end.

### Target prediction

miRNA regulates expression of specific gene via hybridization to mRNA transcripts to promote RNA degradation, inhibit translation or both. For the better understanding of the biological functions of the newly identified Jatropha miRNAs, we have searched for putative target genes using the psRNATarget program with default parameters against the \textit{Arabidopsis thaliana} DFCI Gene Index (AGI) Release 15. We have adjusted expectation value 2.0 for lower false positive prediction. A total of 78 potential targets were identified for the 3 predicted miRNA families which include 21 miRNAs based on their perfect or nearly perfect complementarity with their target sequences in \textit{Arabidopsis} (Table 2). The two miRNA families miR1096 and miR5368 contains 1 and 3 miRNAs does not shows any complementarity with the model plants.

These potential miRNA targets were belonged to a number of gene families that involved in different biological functions such as regulation of metabolism, transcription factor, hormone biosynthesis, development and in oil synthesis (Table 3). The miRNA family ‘miR5021’ showed the highest 51 numbers of independent target genes followed by ‘miR166’ family with 21 numbers of
target genes and miR167 with 6 targets. The rest miRNA families miR1096 and miR5368 does not report any target in the model organism Arabidopsis thaliana within our filtration strategy.

Phylogenetic analysis of predicted miRNAs. Plant miRNAs shows highly conserved nature between distantly related plant species, both at the level of pri-miRNA and mature miRNAs.16 Multiple sequence alignment of the precursor sequences of the predicted miRNAs with other members of the same family showed the high degree of sequence similarity with others for example, the precursor sequence similarity between jcu-MIR166 and other MIR167 members was over 60% (Data not shown). Based on the pre-miRNA sequence comparisons, the evolutionary relationships of J. curcas miRNAs with other members from the same families were analyzed using the ClustalW. It could be seen from the phylogenetic trees that in different families, the evolutionary relationships of J. curcas miRNAs with other species were different; for example, in MIR166 family, jcu-MIR166 and mes-MIR166 were on the same branch (Fig. 3A). Also, it could be seen that different J. curcas miRNA members in the same family were often distantly related (Fig. 3B). These results suggested that different miRNAs might evolve at different rates not only within the same plant species, but also in different ones. Three of five novel miRNAs, namely jcu-miR1096, jcu-miR5368 and jcu-miR5021, for which their homologous miRNAs were not found, showed an unrelated evolutionary relationship with other miRNAs.

**Table 1. Details of the predicted miRNAs from EST**

| miRNAs     | Mature sequence                  | EST      | Loc | NM | LM | LP | G+C | MFE | MFEi |
|------------|----------------------------------|----------|-----|----|----|----|-----|-----|------|
| jcu-miR166a | UCG GAC CAG GCC UCA UUC CCC      | GW880030 | 3'   | 1  | 21 | 146| 39.04| 172.4|3.024646|
| jcu-miR166b | UCG GAC CAG GCC UCA UUC CCC      | GW880030 | 3'   | 2  | 21 | 146| 39.04| 172.4|3.024646|
| jcu-miR166c | UCG GAC CAG GCC UCA UUC CCC      | GW880030 | 3'   | 0  | 21 | 146| 39.04| 172.4|3.024646|
| jcu-miR166d | GUC GGA CCA GCC UUC AUU CCC      | GW880030 | 3'   | 2  | 21 | 144| 38.19| 172.4|3.13491|
| jcu-miR166e | GGA AUG UUG UCC GCU UGG AGG A    | GW880030 | 5'   | 0  | 22 | 142| 37.32| 172.4|3.253174|
| jcu-miR166f | CGU CGG ACC AGG CUU CAU UCC      | GW880030 | 3'   | 2  | 21 | 142| 37.32| 172.4|3.253174|
| jcu-miR166 g | UCG GAC CAG GCC UCA UUC CCC      | GW880030 | 3'   | 1  | 21 | 146| 39.04| 172.4|3.024646|
| jcu-miR166 h | UCG GAC CAG GCC UCA UUC CCC      | GW880030 | 3'   | 0  | 21 | 144| 38.19| 172.4|3.13491|
| jcu-miR167 a | UGA AGC UGC CAG CAU GAU CUG      | GW879969 | 5'   | 1  | 21 | 90  | 44.44| 76.30|1.907691|
| jcu-miR167 b | UGA AGC UGC CAG CAU GAU CUU      | GW881255 | 5'   | 1  | 21 | 68  | 45.59| 129.3|4.170806|
| jcu-miR167 c | UGA AGC UGC CAG CAU GAU CUG A    | GW879969 | 5'   | 1  | 22 | 89  | 43.82| 76.30|1.905642|
| jcu-miR167 d | UGA AGC UGC CAG CAU GAU CUU      | GW881255 | 5'   | 2  | 22 | 68  | 45.59| 129.3|4.170806|
| jcu-miR167 e | AAG CUG CCA GCA UGA UCU A        | GW881255 | 5'   | 1  | 20 | 64  | 45.31| 129.3|4.458867|
| jcu-miR167 f | UGA AGC UGC CAG CAU GAU CUU      | GW881255 | 5'   | 0  | 21 | 67  | 46.27| 129.3|4.170847|
| jcu-miR167 g | UGA AGC UGC CAG CAU GAU CUG A    | GW879969 | 5'   | 1  | 21 | 90  | 44.44| 76.30|1.907691|
| jcu-miR167 h | AUA UCA UGU GGC AGU UUC ACC      | GW879969 | 3'   | 1  | 21 | 93  | 45.16| 76.30|1.816719|
| jcu-miR167 i | UGA AGC UGC CAG CAU GAU CUU      | GW881255 | 5'   | 1  | 21 | 67  | 46.27| 129.3|4.170847|
| jcu-miR1096 | CUG CGU CUU UGG UUC AGG ACU      | GT90456  | 3'   | 2  | 21 | 356 | 42.13| 201.94|3.1346421|
| jcu-miR5368 a | GGA CAG UCU CAG GUA GAC A        | FM889530 | 3'   | 0  | 19 | 166 | 57.23| 163.2|1.717862|
| jcu-miR5368 b | GGA CAG UCU CAG GUA GAC A        | GT95674  | 3'   | 0  | 19 | 166 | 57.23| 83.60|0.879983|

Discussion

Most mature miRNAs are evolutionarily conserved from species to species within the plant kingdom. Therefore, we have used all the previously known plant mature miRNAs from miR registry to search for homology of miRNAs of J. curcas in the publicly available EST database. By computational predictions, we found 20 four miRNAs belonging to five miRNA families. Formation of stem-loop hairpin secondary structure is a critical step in miRNA maturation and is one of the most important characteristics of pre-miRNAs. However, a stem-loop hairpin structure is not a unique feature of miRNAs, because other RNAs (such as mRNA, rRNA, and tRNA) can also form similar hairpin structures. To avoid designating other RNAs or RNA fragments as new miRNAs, several labs have established uniform systems for annotating new miRNAs. Three criteria, such as negative minimal fold energy (MFE), adjusted minimal fold energy (AMFE) and the minimal fold energy index (MFEI) have been proposed. It is indicated that most miRNA precursors identified have an MFEI greater than 0.85, a commonly used value to distinguish miRNAs from other non-coding and coding RNAs. MFE is an important characteristic that determines the secondary structure of nucleic acids (DNA and RNA). Lower the value of the MFE, the higher the thermodynamically stable secondary structure of the corresponding sequence.46 All the mature sequences of Jatropha miRNAs are in the stem portion of the hairpin structures, as shown in Figure 1.

To understand the biological function of miRNAs in plant development, it is necessary to identify their targets. No high-throughput experimental techniques for target site identification have been reported yet. Two strategies have been employed toward this end: (1) genetic approach, which is based on the abnormal expression of target mRNAs in the miRNA loss-of-function
In this paper, with a comparative genomics approach, 24 miRNAs were identified from the EST and GSS databases of *J. curcas*. The findings from this study will contribute to further researches of miRNAs function and regulatory mechanisms.

**Materials and Methods**

**miRNA reference data set.** To search potential miRNAs, the initial miRNA data set has been retrieved from the previously deposited 4677 miRNA of the group viridiplantae from the publically available miRNA database miRBase, version 18.0.\(^1\)

To avoid the overlapping of miRNAs, the repeated sequences of miRNAs were removed with the Jalview program with threshold value to 100.\(^2\) The size of the data set has been subsequently reduced to 2502 non-redundant sequences. These miRNAs were defined as a reference set of miRNA sequences.

**Bioinformatics tools.** Local similarity searches were performed by Blast-2.2.26+ program downloaded from the NCBI ftp site (ftp://ftp.ncbi.nih.gov/). miRNA precursor folding was performed by mFOLD Web server (version 2.3).\(^3\)

**Jatropha EST and GSS.** Total 46862 EST and 1569 GSS (Genomic Survey Sequences) of *J. curcas* were retrieved from the EST and GSS division of GenBank nucleotide database

In this paper, with a comparative genomics approach, 24 miRNAs were identified from the EST and GSS databases of *J. curcas*. The findings from this study will contribute to further researches of miRNAs function and regulatory mechanisms.
The known plant miRNA sequences were subjected to the BLAST search for Jatropha homologs of miRNAs against EST respectively. And all of these sequences were screened against the known plant miRNAs.

**Prediction of potential miRNAs.** There are two important parameters to identify miRNA sequence from the EST and GSS analysis; first is conservation of sequences, and the second is the hairpin stem-loop structure of the potential pre-miRNAs (Fig. 4).

The known plant miRNA sequences were subjected to the BLAST search for Jatropha homologs of miRNAs against EST
and GSS databases. The initial BLAST search was performed with the program of BLAST-2.2.26+, by adjusting the BLASTN parameter settings as: expect values at 1e-3; low complexity was chosen as the sequence filter; the number of descriptions and alignments was raised to 1,000. The default word-match size between the query and database sequences was 7. To be the potent miRNA candidate, the RNA sequences should follow the given criteria:

1. The candidate miRNA should contain at least 18 nt length and does not include any gap.

2. Up to 0–2 nt mismatches in sequence with all previously known plant mature miRNAs were allowed.

The ESTs that closely matched the previously known plant mature miRNAs were included in the set of miRNA candidates and used for additional characterization based on the following criteria:

1. The entire EST sequence was selected to predict the secondary structures and to screen for miRNA precursor sequences.

2. The selected ESTs were further compared with each other to

---

**Figure 2.** (A) The newly identified miRNAs with the different number and (B) miRNA family with its number of miRNA.
eliminate redundancies (3) These precursor sequences were used for BLASTX analysis for removing the protein-coding sequences and retained only the non-protein sequences.

**Prediction of secondary structure.** Precursor sequences of these potential miRNA homologs were used for secondary structure predictions using the Zuker folding algorithm with MFOLD 3.1,13 which is publicly available at www.bioinfo.rpi.edu/applications/mfold/old/rna/. The following parameters were used in predicting the secondary structures:

1. Linear RNA sequence;
2. Folding temperatures fixed at 37°C; ionic conditions of 1M NaCl and with no divalent ions;
3. Percent suboptimility number of 5;
4. Maximum interior/bulge loop size of 30;
5. The grid lines in energy dot plot turned on.

All other parameters were set with default values.

Following criteria were used to choose the candidate miRNA as described by Zhang et al.:15

1. pre-miRNA sequence can fold into an appropriate stem-loop hairpin secondary structure
2. it contains ~22 nt mature miRNA sequence within one arm of the hairpin
3. an MFEI of greater than 0.8536
4. 30–70% A+U content
5. Predicted mature miRNAs had no more than six mismatches with the opposite miRNA* sequence in the other arm
6. Maximum size of 3 nt for a bulge in the miRNA sequence; and
7. No loop or break in miRNA sequences was allowed. These criteria significantly reduced false positives and required that the predicted miRNAs fit the criteria.

ΔG values (kcal/mol) of stem-loop structures generated by MFOLD program were applied to calculate their negative minimal free energies (MFEs), which is directly correlated with the sequence length, to normalize the potential effect of sequence length on MFE and to differentiate miRNAs from other RNAs, we used two energy measurements namely adjusted minimal folding energy (AMFE) and minimal folding free energy index (MFEI). AMFE is defined as the MFE of a 100 nucleotide length.

\[
AMFE = \frac{MFE}{\text{Length of precursor sequence (LP)}} \times 100
\]

The minimal folding free energy index (MFEI) for each sequence was calculated as described by Zhang et al.36

\[
MFEI = \frac{AMFE}{(G + C)\%}
\]

**Prediction of potential targets of miRNAs.** The perfect or near-perfect complementary of miRNAs and their target miRNAs in plants has greatly simplified the identification of miRNA targets. In this study, we applied this strategy to search for the targets of identified miRNAs by homology algorithm. As for *J. curcas*, since only few gene sequences are available, we have used Arabidopsis as a reference system for finding the targets of the candidate miRNAs. The predicted *J. curcas* miRNAs were used as query against the *Arabidopsis thaliana* by using psRNA Target: A Plant small RNA Target Analysis Server to predict the targets of miRNA. http://psRNA.org/psRNATarget/

It should follow given criteria:-

1. Range of central mismatch for translational inhibition 9–11 nucleotide;
2. Maximum expectation value 3
3. Maximum mismatches at the complementary site ≤ 4 without any gaps.
4. Multiplicity of target sites 2

By the parameter set as with default parameters; maximum expectation: 2.0, length for complementarity scoring (hspsize): 20, target accessibility-allowed maximum energy to unpair the target site (UPE): 25.0, Flanking length around target site for target accessibility analysis: 17 bp in upstream and 13 bp in downstream, Range of central mismatch leading to translation inhibition: 9–11nt. The methodology of potential target prediction of predicted miRNA is shown as in Figure 5.

**Phylogenetic analysis.** A homology search of predicted pre-miRNA was done against all plant miRNAs using NCBI standalone BLAST allowing maximum of 3 mismatches and e-value < 0.001. The corresponding precursor sequences of homolog pre-miRNA's were identified and collected. The collected sequences of diverse plant species were aligned with homolog jatropha miRNA using ClustalW. A query of jatropha pre-miRNA against known miRNA's were identified and collected. The collected sequences of miRNAs in plants has greatly simplified the identification of miRNA families. The predicted *J. curcas* miRNAs were adopted in accordance with miRBase.4 The mature sequences are designated ‘MIR’, and the precursor hairpins are labeled as ‘mir’ with the prefix ‘jcu’ for *J. curcas*.

---

**Table 2. Details of the predicted miRNAs from GSS**

| miRNAs    | Mature sequence | GSS       | Loc | NM | LM | LP | G+C | MFE | MFEI |
|------------|-----------------|-----------|-----|----|----|----|-----|-----|------|
| jcu-miR166i| UCGGACCCAGCUCAUCUCGG | JM428509  | 5'  | 0  | 21 | 146| 39.0| 100 | 1.75443 |
| jcu-miR5021a| AGAGAGAAGAAGAAGAAAG | JM428710  | 5'  | 2  | 20 | 68 | 45.6| 84.4| 2.72247 |
| jcu-miR5021b| AGAGAGAAGAAGAAGAAAGAC | JM428982 | 5'  | 2  | 20 | 89 | 43.8| 80.3| 2.05898 |
| Jcu-miR5021c| AAAGAAGAAGAAGAAGAAAG | HN339391  | 5'  | 3  | 20 | 68 | 45.6| 71.1| 2.29346 |

NM: number of mismatch; LM: length of mature miRNAs; LP: length of precursor; MFEs: minimal folding free energies; MFEIs: minimal folding free energy indexes.
| miRNA Acc. | Target Acc. | (E) | Target Description | Target Function |
|------------|-------------|-----|--------------------|-----------------|
| Jcu-miR5021c | TC372725 | 0.0 | hydrolase, acting on glycosyl bonds | Metabolism |
| Jcu-miR5021c | TC360345 | 1.0 | protein kinase | Metabolism |
| Jcu-miR5021c | TC361668 | 1.0 | FIN19.20 | Metabolism |
| Jcu-miR5021c | TC399898 | 1.0 | transferase, transferring pentosyl groups; | Metabolism |
| Jcu-miR5021c | TC358199 | 1.0 | E3 ubiquitin-protein ligase | Metabolism |
| Jcu-miR5021c | TC359005 | 1.0 | E3 ubiquitin-protein ligase UPL3 | Metabolism |
| Jcu-miR5021c | TC369235 | 1.0 | Heparanase-like protein 1 precursor; | Metabolism |
| Jcu-miR5021c | TC364464 | 1.0 | T8K14.20 protein | Biofuel |
| Jcu-miR5021c | TC358567 | 1.0 | phytochrome A supressor spa1 | Metabolism |
| Jcu-miR5021c | NP1660479 | 1.0 | ceramidase family protein | Biofuel |
| Jcu-miR5021c | TC403684 | 1.0 | Receptor-kinase isolog | Metabolism |
| Jcu-miR5021c | TC361202 | 1.0 | ARP protein | Constitutive and alternative splicing |
| Jcu-miR5021c | TC369816 | 1.0 | Calmodulin-binding transcription activator | Signal transduction |
| jcu-miR166a | TC359261 | 2.0 | homeodomain transcription factor | Transcription Factor |
| jcu-miR166a | TC392453 | 2.0 | Class III HD-Zip protein | Transcription Factor |
| jcu-miR166a | TC393398 | 2.0 | HD-Zip protein | Transcription Factor |
| jcu-miR166b | TC359261 | 2.0 | homeodomain transcription factor | Transcription Factor |
| jcu-miR166b | TC392453 | 2.0 | Class III HD-Zip protein 3; | Transcription Factor |
| jcu-miR166c | TC393398 | 2.0 | HD-Zip protein | Transcription Factor |
| jcu-miR166c | TC392926 | 2.0 | homeodomain transcription factor | Transcription Factor |
| jcu-miR166d | TC392453 | 2.0 | Class III HD-Zip protein 3 | Transcription Factor |
| jcu-miR166d | TC393398 | 2.0 | HD-Zip protein | Transcription Factor |
| jcu-miR166d | TC392926 | 2.0 | homeodomain transcription factor | Transcription Factor |
| jcu-miR166d | TC392453 | 2.0 | Class III HD-Zip protein 3 | Transcription Factor |
| jcu-miR166g | TC393926 | 2.0 | homeodomain transcription factor | Transcription Factor |
| jcu-miR166g | TC392453 | 2.0 | Class III HD-Zip protein 3 | Transcription Factor |
| jcu-miR166g | TC393926 | 2.0 | HD-Zip protein | Transcription Factor |
| jcu-miR166h | TC392453 | 2.0 | homeodomain transcription factor | Transcription Factor |
| jcu-miR166h | TC393398 | 2.0 | HD-Zip protein | Transcription Factor |
| jcu-miR166i | TC392926 | 2.0 | homeodomain transcription factor | Transcription Factor |
| jcu-miR166i | TC392453 | 2.0 | Class III HD-Zip protein 3 | Transcription Factor |
| jcu-miR166i | TC393926 | 2.0 | HD-Zip protein | Transcription Factor |
| jcu-miR166i | TC359604 | 1.5 | ARF2 | Hormone Biosynthesis |
| jcu-miR167e | NP230219 | 2.0 | auxin responsive transcription factor | Transcription Factor |
| jcu-miR167e | TC360610 | 2.0 | Auxin response factor 8 | Hormone Biosynthesis |
| jcu-miR167e | TC385699 | 2.0 | Auxin response factor 8 | Hormone Biosynthesis |
| jcu-miR167h | TC361447 | 2.0 | CDPK-related protein kinase | Metabolism |
| jcu-miR167h | TC386736 | 2.0 | CDPK-related protein kinase | Metabolism |
| jcu-miR5021a | NP1662463 | 0.0 | NADK1 (NAD kinase 1) | Metabolism |
| jcu-miR5021a | BX839123 | 0.0 | NAD(H) kinase 1 | Metabolism |
| jcu-miR5021a | TC363441 | 0.0 | NAD(H) kinase 1 | Metabolism |
| jcu-miR5021a | TC379625 | 0.0 | Cytotoxic protein ccdB | Cytotoxic |
| jcu-miR5021a | TC362765 | 0.0 | Diacylglycerol O-acyltransferase | Metabolism |
Table 3. Predicted miRNA targets of Identified miRNAs

| miRNA          | Gene Symbol | Gene Description                                      | Function          |
|----------------|-------------|--------------------------------------------------------|-------------------|
| jcu-miR5021a  | TC370375    | TIF15:13 protein                                       | Metabolism        |
| jcu-miR5021a  | TC358471    | Copia-type polyprotein                                 | Metabolism        |
| jcu-miR5021a  | TC363769    | Copia-type polyprotein                                 | Metabolism        |
| jcu-miR5021a  | TC372725    | Hydrolase, acting on glycosyl bonds                     | Metabolism        |
| jcu-miR5021a  | TC372726    | At1g49510 [Arabidopsis thaliana]                        | Development       |
| jcu-miR5021a  | TC369912    | MAPK3-like protein kinase                              | Metabolism        |
| jcu-miR5021a  | TC363616    | Fucosyltransferase 3                                   | Metabolism        |
| jcu-miR5021a  | TC386702    | pantothenate kinase-related                            | Metabolism        |
| jcu-miR5021a  | NP2693135   | Zeaxanthin epoxidase                                   | Hormone Biosynthesis |
| jcu-miR5021a  | TC361228    | AtABA1 protein                                         | Hormone Biosynthesis |
| jcu-miR5021a  | TC361686    | EMBRYO DEFECTIVE 2738; GTP binding                     | Development       |
| jcu-miR5021a  | TC359684    | At2g41900/T6D20.20                                     | Metabolism        |
| jcu-miR5021a  | TC390065    | (Homeobox-1); transcription factor                    | Transcription Factor |
| jcu-miR5021a  | TC358199    | E3 ubiquitin-protein ligase UPL3                       | Metabolism        |
| jcu-miR5021a  | TC382579    | E3 ubiquitin-protein ligase UPL3                       | Metabolism        |
| jcu-miR5021b  | BX837986    | HB-1 (homeobox-1); transcription factor               | Transcription Factor |
| jcu-miR5021b  | TC390065    | HB-1 (homeobox-1); transcription factor               | Metabolism        |
| jcu-miR5021b  | NP1652463   | NP_188744.3 NADK1 (NAD kinase 1)                       | Metabolism        |
| jcu-miR5021b  | TC363441    | NAD(H) kinase 1                                        | Metabolism        |
| jcu-miR5021b  | TC362765    | Diacylglycerol O-acetyltransferase                     | Metabolism        |
| jcu-miR5021b  | TC370375    | TIF15:13 protein                                       | Growth            |
| jcu-miR5021b  | TC368480    | Br FatA1                                               | Biofuel           |
| jcu-miR5021b  | TC366757    | IAA-amino acid hydrolase ILR1-like 4 precursor        | Hormone Biosynthesis |
| jcu-miR5021b  | NP2693135   | (ABA DEFICIENT 1); zeaxanthin epoxidase               | Hormone Biosynthesis |
| jcu-miR5021b  | TC361228    | AtABA1 protein                                         | Hormone Biosynthesis |
| jcu-miR5021b  | TC361686    | EMBRYO DEFECTIVE 2738; GTP binding                     | Metabolism        |
| jcu-miR5021b  | TC359684    | At2g41900/T6D20.20                                     | Metabolism        |
| jcu-miR5021b  | TC369912    | MAPK3-like protein kinase                              | Metabolism        |
| jcu-miR5021b  | TC358471    | Copia-type polyprotein                                 | Metabolism        |
| jcu-miR5021b  | TC363769    | Copia-type polyprotein                                 | Metabolism        |
| jcu-miR5021b  | TC359534    | replication factor C large subunit-like protein        | Transcription Factor |

Conflicts of Interest
No potential conflicts of interest were found.

Acknowledgments
We are grateful of P.P. Tyag Vallabh Swamiji for his continuous care and blessings. The author gratefully acknowledges the support of Ms. Shivani Patel, Head, Department of Biotechnology, Shree M. and N. Virani Science College, Rajkot and the literature help of Mrs. Sheetal Tank, Librarian, Atmiya Institute of Technology and Sciences. We also wish to thank Visha Rathod, Dr. Priyank Shukla, Nilkanth Faldu, Ravi Ranjan, and Praveen Kumar for their support whenever needed.

References
1. Pavlidis P, Weston J, Cai J, Grundy WN. Gene functional classification from heterogeneous data. In: Proceedings of the fifth annual international conference on Computational biology - RECOMB ’01. New York, New York, USA: ACM Press; 2001.249–255.
2. Lee RC, Feinbaum RL, Ambros V. The C. elegans het erochronic gene lin-4 encodes small RNAs with anti-sense complementarity to lin-14. Cell 1993; 75:843-54; http://dx.doi.org/10.1016/0092-8674(93)90529-Y.
3. Reinhart BJ, Weinstein EG, Rhoades MW, Bartel B, Bartel DP. MicroRNAs in plants. Genes Dev 2002; 16:1616-26; http://dx.doi.org/10.1101/gad.1004402.
4. Griffiths-Jones S, Grocock RJ, van Dongen S, Bareman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006; 34(Database issue):D140-4; http://dx.doi.org/10.1093/nar/gkj112.
5. Jones-Rhoades MW, Bartel DP, Bartel B. MicroRNAs and their regulatory roles in plants. Annu Rev Plant Biol 2006; 57:19-53; http://dx.doi.org/10.1146/annurev.arplant.57.032905.105218.
6. Lee Y, Kim M, Han J, Yeon KH, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. EMBO J 2004; 23:4051-60; http://dx.doi.org/10.1038/sj.emboj.7600385.
7. Park MY, Wu G, Gonzalez-Suarez A, Vaucheret H, Poethig RS. Nuclear processing and export of microRNAs in Arabidopsis. Proc Natl Acad Sci USA 2005; 102:3691-6; http://dx.doi.org/10.1073/pnas.0405570102.
8. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009; 136:215-33; http://dx.doi.org/10.1016/j.cell.2009.01.002.
9. Voinnet O. Origin, biogenesis, and activity of plant microRNAs. Cell 2009; 136:669-87; http://dx.doi.org/10.1016/j.cell.2009.01.046.
10. Zhang B, Pan X, Cobb GP, Anderson TA. Plant microRNA: a small regulatory molecule with big impact. Dev Biol 2006; 295:3-16; http://dx.doi.org/10.1016/j.ydbio.2005.10.036.
Figure 3. Phylogenetic analysis of pre-miRNAs sequences in different families. (A) MIR166 (B) MIR167.
Figure 3. Phylogenetic analysis of pre-miRNAs sequences in different families. (A) MIR166 (B) MIR 167.
32. Pamidimarri DVNS, Sarkar R, Boricha G, Reddy MP. A simplified method for extraction of high quality genomic DNA from Jatropha curcas for genetic diversity and molecular marker studies. Indian J Biotechnol 2009; 8:187-92.
33. Dehgan B. Phylogenetic Significance of Interspecific Hybridization in Jatropha (Euphorbiaceae). Syst Bot 1984; 9:467-78; http://dx.doi.org/10.2307/2418796.
34. Wang CM, Liu P, Sun F, Li L, Liu P, Ye J, et al. Isolation and identification of microRNAs in Jatropha curcas. Int J Biol Sci 2012; 8:418-29; http://dx.doi.org/10.7150/ijbs.3676.
35. Desulian T, Remnert M, Palarnik JF, Weigel D, Huson DH. Identification of plant microRNA homologs. Bioinformatics 2006; 22:359-60; http://dx.doi.org/10.1093/bioinformatics/bti802.
36. Zhang BH, Pan XP, Cox SB, Cobb GP, Anderson TA. Evidence that microRNAs are different from other RNAs. Cell Mol Life Sci 2006; 63:246-54; http://dx.doi.org/10.1007/s00018-005-5467-7.
37. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet 2010; 11:597-610.
38. Dai X, Zhao PX. psRNATarget: a plant small RNA target analysis server. Nucleic Acids Res 2011; 39(Web Server issue):W155-9; PMID:21622958; http://dx.doi.org/10.1093/nar/gkr319.
39. Zhang Y. miRU: an automated plant miRNA target prediction server. Nucleic Acids Res 2005; 33(Web Server issue):W701-4; PMID:15980567; http://dx.doi.org/10.1093/nar/gki383.
40. Das A, Mondal TK. Computational Identification of Conserved microRNAs and Their Targets in Tea (Camellia sinensis). Online 2010; 2010:77-86.
41. Koonmara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 2011; 39(Database issue):D152-7; http://dx.doi.org/10.1093/nar/gkq1027.
42. Clamp M, Cuff J, Searle SM, Barton GJ. The Jalview Java alignment editor. Bioinformatics 2004; 20:426-7; http://dx.doi.org/10.1093/bioinformatics/btg430.
43. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 2003; 31:3406-15; http://dx.doi.org/10.1093/nar/gkg595.
44. Benson DA, Karsch-mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL, et al. GenBank 2002; 30:17-20.
45. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X, et al. A uniform system for microRNA annotation. RNA 2003; 9:277-9; http://dx.doi.org/10.1261/rna.2183803.

Figure 4. Depiction of the steps followed to search for potential miRNAs in J. curcas.
Figure 5. Procedure of potential target search by psRNATarget Server. (http://plantgrn.noble.org/psRNATarget/).

- Identified miRNA sequences in Jatropha curcas
- Arabidopsis thaliana DFCL gene Index (AGI) Database
- Target prediction by psRNA Target
- Functional prediction of potential targets