Analysis of microRNAs and their targets from onion (*Allium cepa*) using genome survey sequences (GSS) and expressed sequence tags (ESTs)

Rukmini Mishra¹, Rupesh Mohapatra², Bijayalaxmi Mahanty³, Raj Kumar Joshi³*

¹School of Applied Sciences, Centurion University of Technology and Management, Odisha, India; ²Centre for Biotechnology, Siksha O Anusandhan University, Bhubaneswar-751030, Odisha, India; ³Dept. of Biotechnology, Rama Devi Women’s University, Vidya Vihar, Bhubaneswar-751022, Odisha, India; Dr. Raj Kumar Joshi - Email: rkjoshi@rdwu.ac.in, *Corresponding author

Received December 27, 2019; Revised December 30, 2019; Accepted December 31, 2019; Published December 31, 2019

DOI: 10.6026/97320630015907

Abstract:
MicroRNAs are small non-coding RNAs of 21-24 nucleotides in length that acts as important modulators of gene expression related to numerous biological processes including development and defense response in eukaryotes. However, only a limited report on onion (*Allium cepa*) miRNAs is available and their associated role in growth and development of onion is not yet clear. Therefore, it is of interest to identify miRNAs and their targets in *Allium cepa* using the genome survey sequences (GSSs) and expressed sequence tags (ESTs) and deduce the functions of the target genes using gene ontology (GO) terms. We report 14 potential miRNAs belonging to 13 different families (miR162, miR168, miR172c, miR172e, miR398, miR400, miR414, miR1134, miR1223, miR6219, miR7725, miR8570, miR8703 and miR8752). BLAST analysis using psRNATarget server predicted 39 potential targets for the identified miRNAs majority of which were transcription factors implicated in plant growth, development, hormone signaling and stress responses. These data forms the basis for further analysis and verification towards understanding the miRNA mediated regulatory mechanism in *Allium cepa*.

Keywords: *Allium cepa*, microRNA, ESTs, GSSs, miRNA targets, gene ontology

Background:
MicroRNAs (miRNAs) are a group of 21-24 nucleotides (nt) small endogenous RNA sequences that acts as negative regulators of gene expression and play significant modulatory roles in numerous biological processes such as growth, development and response to biotic and abiotic stresses [1]. These are basically transcribed out from the endogenous *MIRNA* genes within the intronic and intergenic regions of the eukaryotic genomes in the form of a stem-loop primary miRNA (Pri-miRNA) structure. The pri-miRNA is processed by Dicer-like 1 (DCL1)/Hyponastic Leaves (SE) into hair-pin pre-miRNA and subsequently diced out mature miRNA: miRNA* duplex that are exported into the cytoplasm by HASTY1 (HST1) protein. In the cytosol, the mature miRNA from the duplex binds with the endonuclease ARGONAUTE (AGO) protein forming the RNA induced silencing complex (RISC) and accomplish the regulation of gene expression through cleavage or translational inhibition of the target transcript [2]. Although there are different small non-coding RNAs in plants, miRNAs are unique in the sense that, (1) they are specifically encoded by *MIRNA* genes, (2) possesses a typical stem-loop structure with negative minimal folding free energy (MFE), (3) have a distinct miRNA* sequence, and (4) and exhibit a high degree of sequence complementarity with their specific targets. While a single miRNA can modulate the expression of multiple genes, several miRNAs may also get tangled in the regulation of a specific gene [3]. As such, identification of complementary targets is fundamental to understand the modulatory roles of miRNAs. During the last couple of decades, notable advancement has been made in exposing the fundamental role of miRNAs in plant growth, hormone signaling, organogenesis, floral differentiation and myriads of stress responses [4].

The mature miRNAs are well-conserved through the plant kingdom [5] making it a significant instrument for the identification of novel miRNAs using homology search based
miRNAs were also functionally annotated to understand their and GSS datasets. Further, the target genes of the identified miRNAs using the ESTs datasets important plant species. A previous study had reported 9 onion starting materia
Biotechnology Information (NCBI ESTs and GSSs deposited in the National Centre for molecular and biochemical mechanisms of onion development, targets in onions could provide novel perceptions i
and abiotic stresses
drought, salinity and biotic stresses with price rise due to several environmental factors such as
vegetable crop cultivated in greater parts of the world. Besides being an ingredient with high food value, onion is also credited vegetable crop, including the expressed sequence tags (ESTs) and genome survey sequences (GSSs). Computational analysis of ESTs and GSSs predicted miRNAs with greater precision in various plant species including Arabidopsis [6], wheat [7], maize [8], sorghum [9] and finger millet [10]. The latest release of miRBase (release 22) database has reported 7385 mature miRNA sequences within the Viridiplantae miRNA dataset including 713 from O. sativa, 756 from Medicago truncatula, 401 from Populus trichocarpa, 321 from Zea mays and 241 from Sorghum bicolor.

Bulb onion (Allium cepa L.) is an economically important plant species including the expressed sequence tags (ESTs) and genome survey sequences (GSSs). Computational analysis of ESTs and GSSs predicted miRNAs with greater precision in various plant species including Arabidopsis [6], wheat [7], maize [8], sorghum [9] and finger millet [10]. The latest release of miRBase (release 22) database has reported 7385 mature miRNA sequences within the Viridiplantae miRNA dataset including 713 from O. sativa, 756 from Medicago truncatula, 401 from Populus trichocarpa, 321 from Zea mays and 241 from Sorghum bicolor.

Bulb onion (Allium cepa L.) is an economically important

comparative genomics approach. Several plant miRNAs have been identified through high throughput computational strategies including direct cloning and next generation deep sequencing. However, in cases where the whole genome sequence is not available, the similarity search using Basic Local Alignment Search Tool (BLASTn) for nucleotide sequences within the highly conserved regions of the pre-miRNAs and mature miRNAs as well as matching of the secondary hairpin structure could be effective criterion for miRNA identification in plant species. This is possible by making use of multiple data source including the expressed sequence tags (ESTs) and genome survey sequences (GSSs). Computational analysis of ESTs and GSSs predicted miRNAs with greater precision in various plant species including Arabidopsis [6], wheat [7], maize [8], sorghum [9] and finger millet [10]. The latest release of miRBase (release 22) database has reported 7385 mature miRNA sequences within the Viridiplantae miRNA dataset including 713 from O. sativa, 756 from Medicago truncatula, 401 from Populus trichocarpa, 321 from Zea mays and 241 from Sorghum bicolor.

Bulb onion (Allium cepa L.) is an economically important

plant species including the expressed sequence tags (ESTs) and genome survey sequences (GSSs). Computational analysis of ESTs and GSSs predicted miRNAs with greater precision in various plant species including Arabidopsis [6], wheat [7], maize [8], sorghum [9] and finger millet [10]. The latest release of miRBase (release 22) database has reported 7385 mature miRNA sequences within the Viridiplantae miRNA dataset including 713 from O. sativa, 756 from Medicago truncatula, 401 from Populus trichocarpa, 321 from Zea mays and 241 from Sorghum bicolor.

Bulb onion (Allium cepa L.) is an economically important

Materials & Methods:

**Sequence database and reference set for miRNA identification:**
All mature miRNA sequences from Viridiplantae group were reclaimed from the miRNA database in miRBase (http://www.mirbase.org/) [15]. All these mature miRNAs were previously resulted from different plant species by initial computational identification followed by validation through different experimental approaches including cloning, sRNA sequencing, northern blotting and qPCR method. Mature miRNAs were made non-redundant by duplication to prevent overlapping of miRNA sequences. Taking all these unique mature miRNA sequences as reference, our target miRNA sequences were identified from onion ESTs and GSSs by homology search method. Publicly available 20204ESTs and 10725 GSSs (as of December, 2019) of onion were downloaded from (NCBI) (www.ncbi.nlm.nih.gov/) by using keyword “Allium cepa”.

**Prediction of A. cepa miRNA:**
Prediction process of putative miRNA from Allium cepa is represented in (Figure 1). Sequences from the locally developed onion EST and GST databases were BLAST searched against the GenBank database (www.ncbi.nlm.nih.gov/genbank) and Rfam database ver 12.0 (www.rfam.xfam.org). The resulted sequences were further analyzed with BLASTx [16] to identify and eliminate the coding sequences. The filtered sequences were used for homology search against the known mature and non-redundant plant miRNAs in miRBase (Release 22; http://www.mirbase.org/search.shtml). Sequence alignment of the ESTs and GSSs against the known miRNAs was retrieved through BLASTn algorithms with a threshold E value of 10, sequence filtration at low complexity and word match size between the query and the database set at 7. Homologous candidate miRNAs were identified based on following parameters: EST/GSS sequences with a miRNA matching region of 18 nucleotides with no gap, and base mismatch between predicted sequences and the known miRNAs should be ≥ 3. Zuker algorithm in the MFOLD program predicted the secondary loop structures of the miRNA precursors [17]. The hairpin structures of the precursors were confirmed using the following criteria: hairpin should have at least 18nt mature miRNA in one arm of the stem loop; 50% of bases should be paired; ≤ 4 nt bulge between miRNA and miRNA*; minimum bulge size of 1 or 2 bases and 1 or less asymmetric bulges within the miRNA/miRNA* 30-70% contents of A + U and high negative MFE and minimal folding free energy index (MFEI) of predicted secondary structure.

Negative MFE value of each potential precursor miRNAs were determined by the MFE values (− kcal/mol) of stem-loop structures, which is directly correlated with the sequence length [5]. MFE of a 100 nucleotide length is represented as adjusted minimal folding free energy (AMFE) and is calculated
as: \[ \text{AMFE} = \left( \frac{\text{MFE}}{\text{length of precursor sequence (LP)}} \right) \times 100. \]

Subsequently, the MFEI was calculated as \[ \text{MFEI} = \frac{\text{AMFE}}{(\text{G + C})\%}. \]

![Diagram](image)

**Figure 1:** Schematic approach for the prediction of miRNAs and their targets from ESTs and GSSs of *Allium cepa.*

**Prediction of potential miRNA targets:**
Presumed targets for candidate onion miRNAs were predicted using the psRNA-Target webserver (http://plantgrn.noble.org/psRNATarget/) [18]. Pairwise sequence similarity analyses was performed by querying the *A. thaliana* DFCI gene index (AGI) release 15, *O. sativa* TIGR genome cDNA OSA1 and *A. cepa* DFCI Gene Index (ONGI) cDNA library sequences with the mature onion miRNAs. The default parameters for prediction of miRNA targets are: miRNA-target mismatches \( \leq 2; \) complementary scoring length of 20; maximum energy allowed to unpair the target is 23; flanking length around the target for accessibility (17bp upstream/13bp downstream); range of central mismatch for translational inhibition was set at 9-10nt; number of target for each miRNAs was set at 10.

**Phylogenetic analysis of onion miRNAs:**
The predicted mature onion miRNAs were subjected to nucleotide research using BLASTn against all plant miRNAs as query with a default setting of 3bp mismatch and threshold E value of 0.001. The homologous miRNA precursor sequences were identified and retrieved from miRBase. Multiple sequence alignment of the identified miRNA precursor sequences along with the collected precursors from other plant species was performed using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) with default parameter and manually adjusted using BioEdit 7.1 software. Phylogenetic analyses of the miRNA sequences were performed using Molecular Evolutionary Genetic Analysis (MEGA v 10.1) package [19]. A neighbour joining (NJ) method with 1000 bootstrapping was performed to develop an uprooted phylogenetic tree.
Functional annotation of the miRNA targets:
The functional aspects of the miRNA targets are crucial to comprehend the range of miRNA regulation in the biochemical and molecular mechanism of plant growth and development. Functional enrichment of the miRNA targets was performed using the Blast2GO v3.0 [20] and further verified using the DeepGO prediction tool with the protein GO classes [21]. Identified target genes were categorized in terms of molecular functions, biological processes and cellular components.

Figure 2: Representative hairpin secondary structures of the predicted onion miRNAs as generated by the MFOLD program. Mature miRNAs are indicated with red lines. (a) ace-miR162, (b) ace-miR168, (c) ace-miR172c, (d) ace-miR172e, (e) ace-miR398, (f) ace-miR400, (g) ace-miR414, (h) ace-miR1134, (i) ace-miR1223, (j) ace-miR6219, (k) ace-miR7725, (l) ace-miR8570, (m) ace-miR8703 and (n) ace-miR8752.
Results & Discussion:

Majority of miRNAs are evolutionarily conserved throughout the plant kingdom and therefore could be readily exploited for detection of miRNAs from the ESTs and GSSs of plant species whose complete genome sequence is not yet available [5]. A homology based comparative genome approach was used in the present study for identification of miRNAs from *A. cepa*, a plant with great culinary and economic significance. A total of 7385
non-redundant eukaryotic miRNA sequences from the clade of Viridiplantae were used as reference query for BLAST search against 20225 ESTs and 10725 GSS sequences of *A. cepa*. BLASTn analysis of 503 assembled contigs and 27190 singletons resulted in 54 non-redundant sequences exhibiting high degree of similarity with the Viridiplantae miRNA reference set. BLASTx search against the NCBI protein database identified 29 sequences with protein coding potential and were subsequently eliminated. The remaining 25 sequences were assessed for their ability to develop secondary structure using MFOLD software. A strict evaluation of the hairpin structure predicted only 14 potential miRNAs (6 from ESTs and 8 from GSSs) demonstrating significant sequence identity with conserved Viridiplantae miRNA (Table 1). A recent study carried out by Kohnehrouz et al. [14] reported the identification of 9 miRNAs and their targets from *A. cepa* ESTs. In contrast, we identified 14 miRNAs from *A. cepa* ESTs and GSSs and predicted their detailed functional properties. Interestingly, all except one miRNA (ace-miR414) were new and completely different from those that were identified earlier. This may be attributed to the fact that, both the EST dataset and the GSS database of *A. cepa* was used for miRNA prediction in the present study. Of the 14 miRNAs, 6 were predicted from EST sequences (2.96% per 10,000 ESTs) and 8 were identified from GSSs (7.45% per 10,000 GSSs). These values are quite significant as compared to miRNA candidates from other plant species [5] and confirmed that EST and GSS analysis could be efficiently used to predict miRNAs in plants. Among the newly identified potential *A. cepa* miRNAs, only ace-miR72 consisted of two members (ace-miR72c and ace-miR72e) while the remaining 12 of them (ace-miR162, ace-miR168, ace-miR398, ace-miR400, ace-miR414, ace-miR1134, ace-miR1223, ace-miR6219, ace-miR7722, ace-miR8570, and ace-miR8752) were represented by single member only.ace-miR1223, miR6219, miR7722, miR8570, and miR8752 from *A. cepa* were uniquely reported earlier in *Phycomysembla patens* [22], *Sorghum bicolor* [9], *Brachypodium distachyon* [23], *Amborella trichopoda* [24] and *Gossypium raimondii* [25]. Kohnehrouz et al. [14] identified three unique miRNAs miR1440, miR2673 and miR5021 from *A. cepa* ESTs. The presence of several unique miRNAs in *A. cepa* suggests that they could be involved in the modulation of specific functions related to growth, development and stress responses.

Sequence analysis of miRNA precursors showed that the length of *A. cepa* pre-miRNA and mature miRNA varied from 89 to 200 and 21 to 24, respectively. Plant miRNA biogenesis have shown that the length of plant pre-miRNAs ranges between 100 nt to 1000nt and are usually longer than the animal pre-miRNAs [26]. Our result corroborated with the previously described miRNAs and their precursors in Arabidopsis [27], rice [28], maize [8] so also in *A. cepa* [14]. The identified *A. cepa* miRNAs and their flanking sequences exhibited the typical hairpin loop structures (Fig. 2). Earlier reports have shown that, miRNA precursors have lower folding free energy than the thermodynamic stability of the hair-pin loop [29]. The newly identified onion miRNA precursors exhibited negative MFE values ranging between -48.71 to -142.30 with an average of -76.79±26.81 Kcal/mol (Table 1 & 3). As the length of the miRNA precursor is directly related to MFE value, often the average folding energy (AMFE) and the minimal folding free energy index (MFEI) were calculated to standardise the potential effect of sequence length and differentiate miRNAs from other non-coding RNAs [5]. Usually, the MFEI of the miRNA is significantly higher (>0.85) than other RNAs including tRNA (0.59), tRNAs (0.64) and miRNAs (0.62-0.66) [5]. The AMFE of the predicted onion miRNAs ranged from -36.31 to -75.28 with an average of -52.8±15.17 Kcal/mol. While the previously identified onion miRNAs had average MFEI of 0.67±0.15, the miRNAs identified in the present study had MFEI between 0.89 to 1.8 (1.25±0.31) clearly suggesting that these are precisely distinct and true miRNAs.

Analysis of the nucleotide composition in the pre-miRNAs revealed that uridine (U) is the most representative nucleotide with an average of 30.46% followed by adenine (26.75%), guanine (22.18%) and cytosine (20.57%) (Table 2 & 3). Twelve out of fourteen identified pre-miRNAs consisted of more than 25% of U as has been reported in the miRNAs of other plant species [8, 30]. It was also observed that 8 out of 14 mature miRNAs had U as the preferred 5’ end nucleotide concuring with precedingstudies carried out in canola [31], garlic [32] and chilli [30]. Also, the (G+C) % and the (A+U) % ranged from 26.41-60.5% and 39.5-

| miRNA   | EST/GSS ID  | Length of EST/GSS | Length of mature miRNA | Position | Strand | No. of mismatches | % GC | MFE | AMFE | MFEI |
|---------|-------------|------------------|------------------------|----------|--------|-------------------|-----|-----|------|------|
| ace-miR168 | ET434314    | 830              | 111                    | 21       | 3'     | 0                 | 54.45 | -56.60 | -50.99 | -1.10 |
| ace-miR172 | CF778190    | 825              | 100                    | 21       | 3'     | 0                 | 47.20 | -54.75 | -49.95 | -1.10 |
| ace-miR2176 | CF492133    | 740              | 97                    | 21       | 3'     | 0                 | 38.07 | -49.10 | -45.88 | -1.17 |
| ace-miR2906 | ET487923    | 823              | 101                    | 21       | 3'     | 0                 | 54.45 | -56.60 | -50.99 | -1.19 |
| ace-miR400  | ET482407    | 694              | 107                    | 21       | 3'     | 0                 | 38.07 | -49.10 | -45.88 | -1.17 |
| ace-miR414  | CF468506    | 614              | 111                    | 21       | 3'     | 0                 | 31.01 | -40.71 | -35.88 | -1.34 |
| ace-miR1134 | CF459015    | 428              | 200                    | 24       | 3'     | 0                 | 60.30 | -142.30 | -71.13 | -1.18 |
| ace-miR1223 | ET436105    | 848              | 127                    | 21       | 3'     | 0                 | 39.37 | -51.30 | -40.39 | -1.05 |
| ace-miR2019 | ET485311    | 1056             | 130                    | 24       | 3'     | 0                 | 35.26 | -47.20 | -36.31 | -1.05 |
| ace-miR7225 | CF452200    | 679              | 134                    | 21       | 3'     | 0                 | 44.02 | -91.60 | -68.36 | -1.55 |
| ace-miR9570 | ET464099    | 851              | 159                    | 22       | 3'     | 0                 | 26.41 | -75.70 | -67.60 | -1.80 |
| ace-miR5033 | ET484489    | 857              | 109                    | 24       | 3'     | 0                 | 48.31 | -67.00 | -57.28 | -1.56 |
| ace-miR9072 | CF482208    | 679              | 142                    | 21       | 3'     | 0                 | 46.29 | -80.70 | -54.75 | -1.18 |

Table 1: List of predicted miRNAs from *Allium cepa* and their secondary structure analysis
73.5% respectively. The (A+U) % is well within the range of 30-70%, which is the gold standard for identification of potential plant miRNAs [26]. All the mature miRNAs were located within the stem of the hairpin loop (Figure 2). While 5 miRNAs (41.6%) were positioned in the 5' end of the secondary structure, the remaining 9 (64.28%) were found in the 3' end. This corroborate with the previous finding that majority of onion miRNAs are located on the 3' arm of the secondary hairpin loop structure [14].

Unrooted neighbor-joining phylogenetic tree was developed from the multiple sequence alignment of the identified onion miRNAs and other members of the same family available in miRBase to determine the evolutionary relationship among them. Distinct phylogenetic trees were obtained for ace-miR162, ace-miR168, ace-miR172c, ace-miR172e and ace-miR398 exhibiting high degree of sequence similarity with miRNAs from other plant species (Figure 3). ace-miR172c demonstrated 27-54% of
miRNA modulate the expression of target mRNA through complementary binding and consequent cleavage and/or translational inhibition [33]. To understand the functional role of the identified onion miRNAs, potential miRNA target genes were predicted using the psRNATarget webserver with default parameters. As the target sites of plant miRNAs are mostly situated in the open reading frames (ORFs) of the target genes, onion ESTs in addition to AGI and TAIR databases were used to search for putative target genes. Based on accepted principles, the algorithm predicted 39 potential target for 13 miRNAs (Table 4).

Unsurprisingly, most of the target were similar to the one that have been earlier validated as plant miRNA targets in other plant species including Arabidopsis, rice, wheat, maize and garlic [32]. Nine miRNAs (ace-miR162, ace-miR168, ace-miR172c, ace-miR172e, ace-miR398, ace-miR398c, ace-miR400, ace-miR414, ace-miR1134 and ace-miR6219) were predicted to have targets in the range of 2 to 8 suggesting that these miRNAs might have diverse functional attributes. Among the 39 targets, 10 were transcription factor (TF) genes including ethylene response factors (ERFs), MADS-Box TFs and Ring-H2 proteins that have been previously implicated in plant growth regulation and development [34]. Most of the targets exhibited high homology with targets from other plants and presumably demonstrated functional redundancy across plant species. For example, ace-miR162 targeting Dicer-like (DCL) proteins, ace-miR168 targeting Argonaute 1 (AGO1) proteins, ace-miR400 targeting pentatrico peptide repeat protein 1 (PPR1) and ace-miR1134 targeting receptor protein kinase PERK1 have been formerly involved in gene regulation and small RNA biogenesis, plant growth, stress response, hormone signalling and host-microbe interactions [31, 32, 35].ace-miR172c targeted two genes encoding floral homeotic protein APETALA2 suggesting its involvement in the speciation of onion flowers. A few targets were non-transcription factors such as phosphoenolpyruvate carboxylase (ace-miR1134), protein phosphatase (ace-miR8752) and ribosome inactivating protein 1 (ace-miR6219) vindicating their functional role related to plant metabolism, immunity and defense response [36]. Additionally, 8 genes targeted by ace-miRNAs were uncharacterized with unknown function, suggesting that they could be part of unknown biochemical and molecular mechanisms essential for growth and survival of the plant.

### Table 4: Putative targets of the identified miRNAs from Allium cepa

| miRNA name | Target Accn. | Targeted McGregor (bp) | Target Description | Expectation | UPE | Inhibition |
|------------|--------------|-------------------------|--------------------|-------------|-----|------------|
| ace-miR162 | LOC_Os05g32180 | CUGAUGAUCAGCUAGUCUUGA | Dicer-like protein (DCL) | 1.0 | 14.1 | Cleavage |
|            | LOC_Os06g27290 | CUGGAGAUCUGCUAGAGUUGA | Dicer-like protein (DCL) | 1.8 | 17.8 | Cleavage |
|            | LOC_Os07g36740 | CUGAUGCAGUCCGCUAGUUGA | Dicer-like protein (DCL) | 1.3 | 17.1 | Cleavage |
|            | LOC_Os06g26640 | CUGCAUGCAUCUAAUGCUAACU | AGO1 protein | 0.3 | 17.1 | Cleavage |
|            | LOC_Os02g51500 | CUGCAUGCAUCUAAUGCUAACU | AGO1 protein | 0.3 | 17.1 | Cleavage |
| ace-miR172e| TC0745742 | CUGAAGAUGGAGACACGUACCU | Uncharacterized protein | 1.5 | 15.9 | Cleavage |
|            | RP053546 | AAGCUUCAUGCAUCUAGCCCU | Uncharacterized protein | 1.8 | 16.5 | Cleavage |
|            | TC097809 | AAUCUCAUGCAUCUAGCCCU | Uncharacterized protein | 1.1 | 14.2 | Cleavage |
|            | TC096720 | UUGAUGCAUCUAGCCUACCU | Uncharacterized protein | 1.0 | 14.9 | Cleavage |
|            | LOC_Os04g55120 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 1.5 | 15.9 | Cleavage |
|            | LOC_Os04g55561 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.8 | 14.0 | Cleavage |
|            | LOC_Os04g55562 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
|            | TC094989 | UUGAUGCAUCUAGCCUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
|            | LOC_Os04g55120 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 1.5 | 15.9 | Cleavage |
|            | LOC_Os04g55561 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
|            | LOC_Os04g55562 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
|            | TC094989 | UUGAUGCAUCUAGCCUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
| ace-miR172e| TC0745742 | CUGAAGAUGGAGACACGUACCU | Uncharacterized protein | 1.5 | 15.9 | Cleavage |
|            | RP053546 | AAGCUUCAUGCAUCUAGCCCU | Uncharacterized protein | 1.8 | 16.5 | Cleavage |
|            | TC097809 | AAUCUCAUGCAUCUAGCCCU | Uncharacterized protein | 1.1 | 14.2 | Cleavage |
|            | TC096720 | UUGAUGCAUCUAGCCUACCU | Uncharacterized protein | 1.0 | 14.9 | Cleavage |
|            | LOC_Os04g55120 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 1.5 | 15.9 | Cleavage |
|            | LOC_Os04g55561 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.8 | 14.0 | Cleavage |
|            | LOC_Os04g55562 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
|            | TC094989 | UUGAUGCAUCUAGCCUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
| ace-miR172e| TC0745742 | CUGAAGAUGGAGACACGUACCU | Uncharacterized protein | 1.5 | 15.9 | Cleavage |
|            | RP053546 | AAGCUUCAUGCAUCUAGCCCU | Uncharacterized protein | 1.8 | 16.5 | Cleavage |
|            | TC097809 | AAUCUCAUGCAUCUAGCCCU | Uncharacterized protein | 1.1 | 14.2 | Cleavage |
|            | TC096720 | UUGAUGCAUCUAGCCUACCU | Uncharacterized protein | 1.0 | 14.9 | Cleavage |
|            | LOC_Os04g55120 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 1.5 | 15.9 | Cleavage |
|            | LOC_Os04g55561 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.8 | 14.0 | Cleavage |
|            | LOC_Os04g55562 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
|            | TC094989 | UUGAUGCAUCUAGCCUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
| ace-miR172e| TC0745742 | CUGAAGAUGGAGACACGUACCU | Uncharacterized protein | 1.5 | 15.9 | Cleavage |
|            | RP053546 | AAGCUUCAUGCAUCUAGCCCU | Uncharacterized protein | 1.8 | 16.5 | Cleavage |
|            | TC097809 | AAUCUCAUGCAUCUAGCCCU | Uncharacterized protein | 1.1 | 14.2 | Cleavage |
|            | TC096720 | UUGAUGCAUCUAGCCUACCU | Uncharacterized protein | 1.0 | 14.9 | Cleavage |
|            | LOC_Os04g55120 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 1.5 | 15.9 | Cleavage |
|            | LOC_Os04g55561 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.8 | 14.0 | Cleavage |
|            | LOC_Os04g55562 | CUGAAGAUGGAGACACGUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
|            | TC094989 | UUGAUGCAUCUAGCCUACCU | Dicer-like protein (DCL) | 0.5 | 13.5 | Cleavage |
To delineate the comprehensive network of genes modulated by miRNAs, the identified targets were subjected to gene ontology (GO) term analysis using the Blast2Go program. A total of 31 out of 39 predicted targets were categorized into 8 biological processes, 5 molecular functions and 4 cellular components (Table 5). Among the biological processes, genes involved in metabolic process (10), secondary metabolic process (5), signaling (5), regulation of transcription (5) and response to stress (4) were mostly represented. Two target genes (CF451171, CF452115) were specifically involved in immune system process (GO: 0002376) and defense response (GO: 0006952). Likewise, transcription factor activity (GO: 0003700; 7 genes), Catalytic activity (GO: 0003676) and nucleic acid binding (GO: 0003824; 7 genes) were the most represented GO terms in the molecular function category. As regards to the putative target transcript of miRNAs in the cellular component category, the GO term cell part (GO: 0044464) was most represented with 5 genes followed by intracellular part (GO: 0044424) and organelle (GO: 0043226) with 3 genes each. The diversified function of these target genes suggest that the complementing miRNAs presumably plays important modulatory role in the signalling, growth, development and defense response to myraids of stresses in onions.

| GO IDs | Description | E-value | genes | Access. ID for the target | miRNAs |
|--------|-------------|---------|-------|--------------------------|---------|
| GO:0003676 | Nuclear acid binding | 2.4E-03 | 6 | LOC_Os04g55560.1, LOC_Os02g08364.1, LOC_Os11g0168100, LOC_Os04g0501000, LOC_Os04g3874000 | miR156, miR172, miR390, miR404 |
| GO:0003513 | Protein binding | 1.1E-03 | 2 | LOC_Os09g14670.1, LOC_Os09g14670.2, LOC_Os05g1648000, LOC_Os09g14670.3 | miR156, miR172, miR390, miR404 |
| GO:0003824 | Catalytic activity | 4.5E-04 | 7 | LOC_Os09g14670.1, LOC_Os09g14670.2, LOC_Os05g1648000, LOC_Os09g14670.3, LOC_Os05g4322000, LOC_Os05g4322000, LOC_Os05g4322000 | miR156, miR172, miR390, miR404, miR8752 |
| GO:0003254 | Enzyme regulator activity | 3.6E-03 | 3 | LOC_Os09g14670.1, LOC_Os09g14670.2, LOC_Os05g1648000, LOC_Os09g14670.3, LOC_Os05g4322000, LOC_Os05g4322000, LOC_Os05g4322000 | miR156, miR172, miR390, miR404, miR8752 |
| GO:0003700 | Cellular component | 2.8E-01 | 7 | LOC_Os09g14670.1, LOC_Os09g14670.2, LOC_Os05g1648000, LOC_Os09g14670.3, LOC_Os05g4322000, LOC_Os05g4322000, LOC_Os05g4322000 | miR156, miR172, miR390, miR404, miR8752 |

Conclusion:
In conclusion, a comprehensive computational analyses of onion ESTs and GSSs were performed in the present study to identify 14 potential miRNAs belonging to 13 different families. Phylogenetic analysis of the identified miRNAs confirmed their close homology with conserved miRNAs from other plant species. A total of 39 potential targets were predicted for the identified miRNAs with an inhibitive expression response due to miRNA mediated cleavage or translational repression. Bulk of the predicted target genes encoded transcription and regulatory factors that are implicated in plant growth, development, hormone signalling and stress responses. GO annotation of the target genes revealed that the miRNAs and their associated components are significant modulators of metabolic processes, plant immunity and defense response. These dataset will form the basis for further characterization of miRNAs through transient over-expression and knockout study towards exploration of miRNA mediated regulatory mechanism in onion.

Acknowledgment:
A part of this study is funded by a research grant (Grant no. BT.PR23412/BPA/118/284/2017) from the Dept. of
Biotechnology, Govt. of India. BM is grateful to Dept. of Biotechnology (DBT), Govt. of India for financial support in form of Junior Research Fellowship. The authors are thankful to the Head, Dept. of Biotechnology, Rama Devi Women’s University, Vidya Vihar, Bhubaneswar for her guidance and support.

References:
[1] Sunkar R et al. Trends Plant Sci. 2012 17: 196. [PMID: 22365280]
[2] Naqvi AR et al. J. Cell Physiol. 2012 227:3163. [PMID: 22252306]
[3] Dehury B et al. Plant Signal. Behav. 2013. 8:e26543. [PMID: 24067297]
[4] Li C and Zhang B J Cell Physiol. 2016 231:303. [PMID: 26248304]
[5] Zhang BH et al. Plant J. 2006 46: 243. [PMID: 16623887]
[6] Wang XJ et al. Genome Biol. 2004 5: R65. [PMID: 15345049]
[7] Jin W et al. J Plant Res. 2008 121:351. [PMID: 18357413]
[8] Zhang L et al. PLoS Genet. 2009 5: e1000716. [PMID: 19936050]
[9] Katiyar A et al. Plant Signal. Behav. 2012 7: 1. [PMID: 22415044]
[10] Usha S et al. InterdiscipSciComput Life Sci. 2017 9:72. [PMID: 26496774]
[11] Smith C et al. Tex J Agric Nat Resour. 2003 16:24.
[12] Cramer CS Euphytica. 2000 115:159.
[13] Nanda S et al. Plant Pathol. J. 2016 32: 519. [PMID: 27904458]
[14] Kohnehrouz BB et al. InterdiscipSciComput Life Sci. 2017 10: 771. [PMID: 28660536]
[15] Griffiths-Jones S et al. Nucleic Acids Res. 2008 36: D154. [PMID: 17991681]

[16] Altschul SF et al. J. Mol. Biol.1990 215: 403. [PMID: 2231712]
[17] Zuker M Nucleic Acids Res. 2003 31:3406. [PMID: 12824337]
[18] Dai X et al. Nucleic Acids Res. 2011 39. W155. [PMID: 21622958]
[19] Kumar S et al. MolBiolEvol. 2018 35:1547. [PMID: 29722887]
[20] Conesa A and Götz S. Int J Plant Genomics. 2000 2008: 619832. [PMID: 18483572]
[21] Kulmanov M et al. Bioinformatics. 2018 34:660. [PMID: 29028931]
[22] Talmor-Neiman M et al. Plant J. 2006 48: 511. [PMID: 17076803]
[23] Bertolini E et al.Mol. Plant. 2013 6:423. [PMID: 23264558]
[24] Amborella Genome project. Science. 2013 342:1241089. [PMID: 24357323]
[25] Xue W et al. BMC Genomics. 2013 14:629. [PMID: 24044642]
[26] Meyers BC et al. Plant Cell 2008 20:3186. [PMID: 19074682]
[27] Sunkar R and Zhu JK Plant Cell 2004 16: 2001. [PMID: 15258262]

License statement: This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License

Articles published in BIOINFORMATION are open for relevant post publication comments and criticisms, which will be published immediately linking to the original article for FREE of cost without open access charges. Comments should be concise, coherent and critical in less than 1000 words.
