International Journal of Pharmaceutical Science Invention
ISSN (Online): 2319 – 6718, ISSN (Print): 2319 – 670X
www.ijpsi.org Volume 2 Issue 5 ǁ May 2013 ǁ PP.70-74

Identification of miRNAs in *Eucalyptus globulus* Plant by Computational Methods

P. Ranjit[1*], Dr. Archana Giri[2], G. Sowjanya[3], M. Vennela[4] , G.Vineela[5]

1,2,3,4 Centre for Biotechnology, Institute of Science and Technology,
Jawaharlal Nehru Technological University, Hyderabad – 500085, A.P., India
[5]Muvva Biosolutions Pvt.Ltd, Bioinformatics Division, #301, No: 7-1-414/21,
Kaushik Sai Rama Residency, Srinivasa colony (East) S.R.Nagar, Hyderabad-500038, A.P, India.

ABSTRACT: *MicroRNAs (miRNAs) are a novel growing family of endogenous, small, non- coding, single stranded RNA molecules directly involved in regulating gene expression at the post transcriptional level. High conservation of miRNAs in plant provides the foundation for identification of new miRNAs in other plant species through homology alignment. Here, previous known plant miRNAs were BLASTed against the Expressed Sequence Tag (EST) database of Eucalyptus globulus, and according to a series of filtering criteria, a total of 7 miRNAs belonging to 3 miRNA families were identified. Overall, our findings lay the foundation for further researches of miRNAs function in E.globulus.*

Keyword*: Eucalyptus globulus, Expressed sequence tag,, MicroRNA, Minimal free energy index***.**

I. INTRODUCTION

MicroRNAs (miRNAs) are a novel growing family of endogenous, small, non- coding, single-stranded RNA molecules encoded in the genomes of plants and animals that repress mRNA translation or mediate mRNA degradation in a sequence-specific manner **[1].** The discovery of the first microRNA *lin-4* in *Caenorhabditis elegans* by Ambros laboratory emerged as biology's unusual or unique findings **[2].** These tiny bits of RNA play a major role in gene regulation, which involves in negative regulation of gene targets. In recent years, identification and functional studies of miRNA has made great progress in research. The exposition of miRNA in plants is still a continuing process hence, till date a number of plant miRNAs have been discovered and functionally identified. Key roles of miRNA in biological processes are revealed by plant studies, which include regulation of leaf development **[3],** stem development **[4],** root development **[5],** signal transduction, developmental timing, various floral differentiation and development, and defense response against stresses. Plant miRNAs are usually evolutionary conserved sequences and are observed in regions of the genome distinct from previously annotated genes.

Biomedical Informatics Different approaches used for miRNA identification includes gene cloning technology and Bioinformatics strategies. Gene cloning is a conventional method to identify the new miRNA accurately, even though it has disadvantages, such as difficulty in finding miRNAs which express at low levels, difficulty in cloning, degradation of RNA during sample separation etc **[6].** Rapid development in the field of bioinformatics has brought a number of computational programs and other tools to successfully predict the miRNA **[7, 8].** This process is purely based on the genomic databases like expressed sequence tags (EST) and other like genome survey sequences (GSS). Since the miRNAs are more conserved in plant species, it is possible to identify novel miRNAs using computational techniques. Now a days miRNAs are identified using the computational or bioinformatics based approach, as it is very useful in predicting the novel miRNA, which cannot be done by cloning. *E. globulus* leaves possess a number of pharmacological activities. E. globolus is a medicinal plant that has been found to possess analgesic, anti bacterial, anti inflammatory properties etc., E. globolus is rich in cineole, pinene and other compounds like phellandrene etc. which makes it a very effective therapeutic plant. Eucalyptus oil has a cooling and deodorising effect on the body and is against fever, migrane and malaria**.**

In this study, all previously known plant miRNAs from A.thaliana, rice, and other plant species were used to search the E. globulus homologs of miRNAs in the publicly available expressed sequence tag (EST), National Center for Biotechnology Information, NCBI, (http://www.ncbi.nlm.nih.gov/).
A total of 7 potential miRNAs were detected. miRNAs were found to be coding transcription factors which are involved in regulating plant growth, development and metabolism.

II. METHODOLOGY

2.1 Datasets of miRNAs, EST and mRNA sequences

To search potential miRNAs in E. globulus, previously known plant miRNAs including their precursor sequences from *Arabidopsis thaliana*, *Zea mays*, *Oryza sativa*, *Glycine max*, *Sorghum bicolor* and other plant species were downloaded from the miRBase (Release 19:August 2012) (http://www.mirbase.org/) **[7].** E. globulus expressed sequence tags (EST) were downloaded from GenBank database (http://www.ncbi.nlm.nih.gov).

2.2 Availability of Computational Software

Comparative software BLAST-2.2.24 [8] was downloaded from NCBI. BLASTX (http://www.ncbi.nlm.nih.gov/BLAST/) was used for analysis of potential targets. RNA secondary structure and the free energy were calculated by web server Mfold 3.2 **[9].**

2.3 Prediction of E. globulus miRNAs

The Computational prediction of miRNAs in E. globulus is displayed in Fig: 1**.** the plant miRNA sequences were initially scanned to remove repeats. The reference set was subjected to the BLAST **[10]** search for E. globulus homologs of miRNAs against EST database. The initial BLAST-2.2.24 search was carried out with default parameters. Mature miRNA sequence should be no less than 18 nt, and the mismatches should be less than 4. Later these precursor sequences were BLASTXed online to remove the protein coding sequences. Pre-miRNAs secondary structure was run on MFOLD 3.2 (http://www.bioinfo.rpi.edu/).

The following steps were considered for screening the candidate miRNA homologs: (1) The RNA sequence folding into an appropriate stem-loop hairpin secondary structure that contains the ~22 nt mature miRNA sequence located in one arm of the hairpin structure; (2) The predicted mature miRNAs with no more than 6 mismatches with the opposite miRNA* sequence in the other arm; (3) maximum size of 3 nt for a bulge in the miRNA sequence was allowed; (4) miRNA precursors with secondary structures had higher negative minimal free energies and minimal free energy index (MFEI) than other different type of RNAs; and MFEI of greater than 0.85 **[11]**; (5) The A+U content of pre-miRNA within 30-70% was considered; (6) no loop or break in miRNA sequences was allowed. These criteria significantly reduced false positives and required that the predicted miRNAs fit the criteria proposed by Ambros and co-workers **[12].**

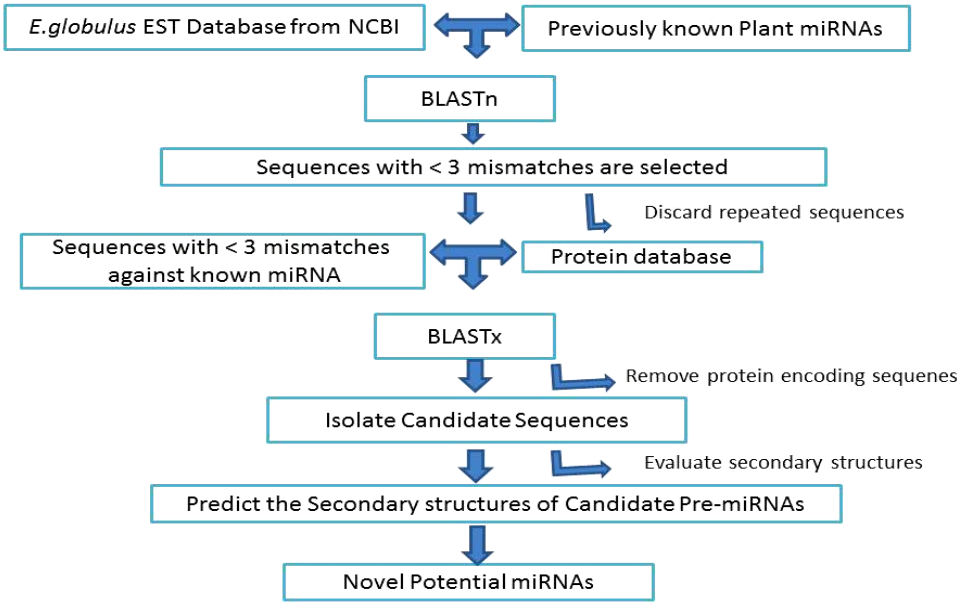

Fig1: Flowchart of *E. globulus* miRNA prediction

2.4 Computational prediction of potential targets of miRNAs

miRNA targets prediction was performed by aligning the predicted miRNA sequences with mRNA sequences of E. globulus through the BLAST program. The targets were screened according to these criteria: the number of mismatches should be less than 4, and no gaps were allowed at the complementary sites. After removal of the repeated sequences, the potential target genes were BLASTed against protein databases to predict their function.

III. RESULTS AND DISCUSSION

To signify new miRNAs in *E. globulus* by Bioinformatics strategy, the flowchart is shown in **Figure.1**. Setting the default e-value, we began our BLASTn search for homologous miRNA sequences against the EST databases of E. globulus .After the blast, all blasted hits with non-coding sequences were retained for analysis of secondary structure; those meeting the criteria, discussed in Methodology were termed as miRNA homologs. Finally, 7 potential E. globulus miRNAs belonging to 3 miRNA families were identified and they were named according to Ambros **[12].** The details on predicted E. globulus miRNAs including family names, Sources, mismatches, mature miRNA lengths, precursors, A+U Content, and Minimal folding free energy index (MFEI) were listed in **Table 1**. During the screening of the potential miRNAs, candidate miRNAs were evaluated for A+U content. The sequences of the miRNA precursors had A+U content ranging from 31.57% to 61.90% **Table 1,** which is in agreement with the previous results **[13].** The length of the 48 predicted miRNAs ranged from 18nt to 22nt. All the MFEI of these hairpin structures were over 0.85, which was thought to be gold standard to differentiate miRNAs from other ones **[14].** The 7 miRNAs represent 3 miRNA families in E. globulus **(Figure. 2).** miRNA 399 has two members; miRNA 169 have four members; miRNA 167 have 1 members. The current results confirm that the approach for EST analysis is a relatively efficient way to identify miRNAs.

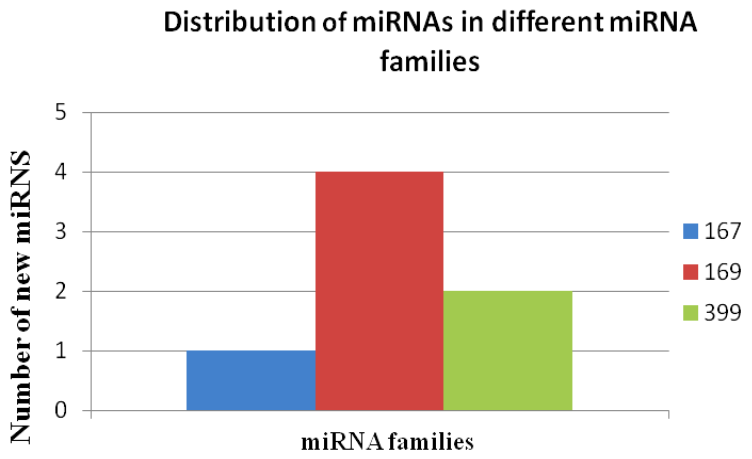

Fig 2: The distribution of new miRNAs in different miRNA families

According to Zhang **[13],** about 10000 ESTs contained one miRNA, so about 3 miRNAs should be predicted theoretically from the total of 29448 ESTs, but in this work we found total 7 miRNAs belonging to 3 families showing the higher value than the previous prediction results and different length of mature miRNAs from the same precursor were regarded as different ones, considering they corresponded to different target genes. Compared to the nucleotide number of animal miRNA precursors (typically with 70–80), the E. globulus miRNA precursors show more diversity in structure and size **Table 1.** The length of miRNA precursors in E. globulus varied from 64 to 69 nucleotides. 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 **[13].** The diversity of the identified miRNAs can also be found in the location of mature miRNA sequences. The sequences of miR 399 a/b, was located at the 5′ end of the miRNA precursors, while the miR 169 a/b/c/d, miR 167 were found at the 3′ end. Based on the complementarity between miRNAs and their target genes in plants, the E. globulus EST database was searched for homology to the new miRNA sequences with a BLASTN and BLASTX algorithm for the discovery of miRNA targets. A total of 2 potential targets for E. globulus miRNAs were identified. These potential miRNA targets belonged to a number of gene families that had different biological.

In plant kingdom most of the mature miRNAs are evolutionarily conserved from species to species. This information enables us to predict new miRNA homologs or orthologs by insilico method **[13].**Therefore, we used all previously known plant mature miRNAs from miRBase to search for homologs of miRNAs and their target genes in radish in the publicly available EST and GSS database of E. globulus . Finally, 7 potential E. *globulus* miRNAs belonging to 3 miRNA families were identified. In the present study, the length of predicted miRNA precursors varies from 64 to 69 NT. The different sizes of the identified miRNAs within the different families suggest that they may perform unique functions in the regulation of miRNA biogenesis or gene expression **[13].** MFEI is an important characteristic that distinguish miRNA from other non-coding and

coding RNAs. The MFEI is a unique criterion to designate miRNAs. When the MFEI is more than 0.85, the sequence is most likely to be miRNA. We observed that the MFEIs of the hairpin structures ranged from 1.54 to 2.08 Table **1**. All the mature sequences of E. *globulus* miRNAs are in the stem portion of the hairpin structures, as shown in Figure 3. According to the estimation approximately 10,000 ESTs in plants contain one miRNA **[13]** that means 29448 ESTs in E.globulus should contain 3 miRNAs. But in this study, 7 miRNAs were detected, showing higher value than the previous prediction **(Figure. 3)**. The current results confirm that the approach of EST analysis is a relatively efficient way to identify miRNAs. To understand the biological function of miRNAs in plant development, it is necessary to identify their targets. In miRNA target prediction, the screening criterion was set according to the description in Methodology. Finally, 2 potential targets for E. globulus miRNAs were identified**.** The miRNA 169 have been predicted to target cytochrome c oxidase subunit 2 gene and miRNA 167 targeted unknown protein name. The general characteristic of the miRNA sequence is, it is complementary to their target gene, and in some case single miRNA can be complementary to more than one target gene. Cytochrome c oxidase subunit 2 is a component of the respiratory chain and is involved in the transfer of electrons from cytochrome c to Oxygen. . In eukaryotes this enzyme complex is located in the mitochondrial inner membrane. The findings of this study considerably broaden the scope of understanding the function of miRNA in E. globules

Table 1: Predicted miRNAs of *E. Globules*

New miRNAs	EST ID	Mature sequence	NM	ML	LP	A+U (%)	MEFI
egl –miR399a	**162328456**	**tgccaaaggagagttgccct**	**1**	**19**	**69**	**45**	**2.08**
egl –miR399b	**162328456**	**tgccaaaggagagttgccctt**	**0**	**20**	**69**	**50**	**2.08**
egl –miR167	**162327474**	**tgaagctgccagcgtgatct**	**1**	**19**	**68**	**45**	**2.23**
egl –miR169a	**162325283**	**agccaaggatgacttgcc**	**0**	**18**	**64**	**44.44**	**1.54**
egl –miR169b	**162325283**	**cagccaaggatgacttgcc**	**0**	**19**	**64**	**42.10**	**1.54**
egl –miR169c	**162325283**	**cagccaaggatgacttgccg**	**0**	**20**	**64**	**40**	**1.54**
egl –miR169d	**162325283**	**cagccaaggatgacttgccga**	**0**	**21**	**64**	**42.85**	**1.54**

NM=Number of Mismatch, ML=Length of Mature miRNAs, LP=Length of Precursor, A+U=(Adenine +uracil)%. MFEI=Minimal free energy index.

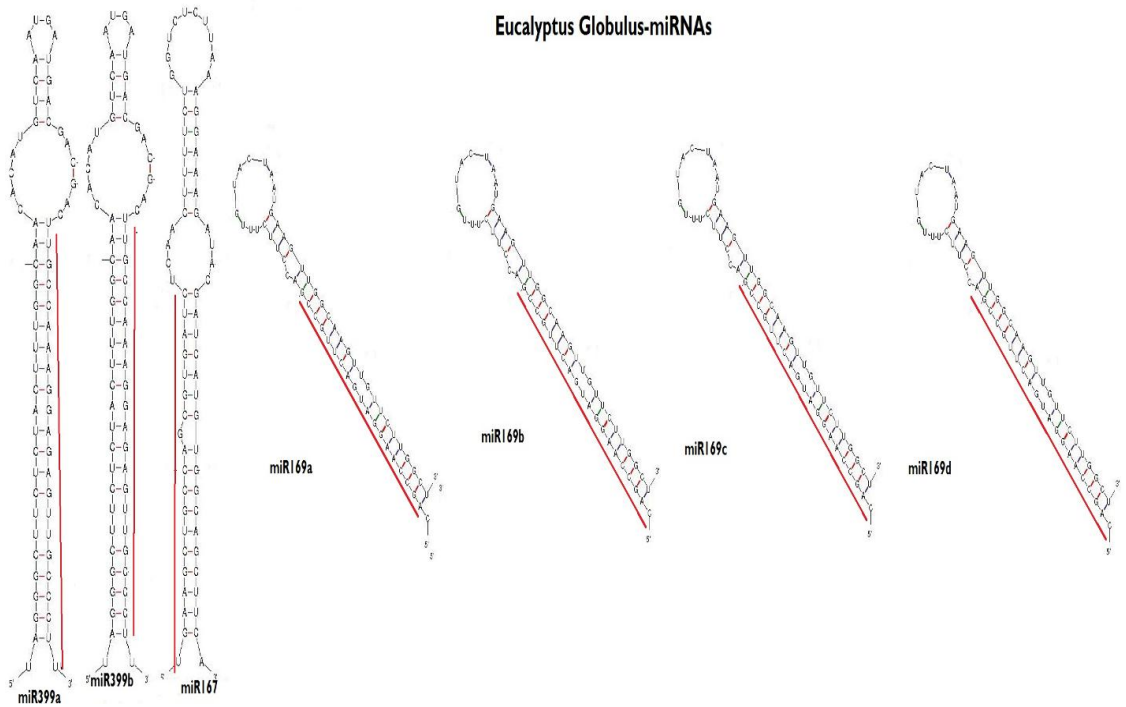

Fig 3: Mature and precursor sequences and the predicted stem-loop structures of newly Identified miRNAs in *E. globulus.* The mature miRNAs are underlined.

VI. CONCLUSION

This paper, with a bioinformatics approach, 7 mature miRNAs along with 2 target genes were identified in E. *globulus*. In-silico studies stand as initial point for understanding miRNAs role in gene regulation. Thus, identification of miRNAs and their target genes help in understanding function and regulatory mechanisms in E. *Globulus.*

REFERENCES

[1]. Unver T et al. Int J Plant Genomics. 2009: 262463 [PMID: 19834623]
[2]. Lee RC et al. Cell. 1993 75: 843 [PMID: 8252621]
[3]. Palatnik JF et al. Nature. 2003 425: 257[PMID: 12931144]
[4]. Mallory AC et al. Curr Biol. 2004 14: 1035[PMID: 15202996]
[5]. Guo HS et al. Plant Cell. 2005 17: 1376 [PMID: 15829603]
[6]. Zhang BH et al. Cell Mol Life Sci. 2006 63: 246 [PMID: 16395542]
[7]. Griffiths-Jones S, Methods Mol Biol. 2006 342: 129[PMID: 16957372]
[8]. Altschul SF et al. Nucleic Acids Res. 1997 25: 3389[PMID: 9254694]
[9]. Zuker M. Nucleic Acids Res. 2003 31: 3406 [PMID: 12824337]
[10]. Mathews DH et al. J Mol Biol. 1999 288: 911 [PMID: 10329189]
[11]. Zhang BH et al. Cell Mol Life Sci. 2006 63: 246[PMID: 16395542]
[12]. Ambros V et al. RNA. 2003 9: 277 [PMID: 12592000]
[13]. Zhang B et al. Plant J. 2006 46: 243 [PMID: 16623887]