**In silico Identification of miRNAs and Their Targets in Cluster Bean for Their Role in Development and Physiological Responses**

Vrantika Chaudhary, Sumit Jangra and Neelam R. Yadav*

Department of Molecular Biology, Biotechnology and Bioinformatics, CCS Haryana Agricultural University, Hisar, India

Cluster bean popularly known as “guar” is a drought-tolerant, annual legume that has recently emerged as an economically important crop, owing to its high protein and gum content. The guar gum has wide range of applications in food, pharma, and mining industries. India is the leading exporter of various cluster bean-based products all across the globe. Non-coding RNAs (miRNAs) are involved in regulating the expression of the target genes leading to variations in the associated pathways or final protein concentrations. The understanding of miRNAs and their associated targets in cluster bean is yet to be used to its full potential. In the present study, cluster bean EST (Expressed Sequence Tags) database was exploited to identify the miRNA and their predicted targets associated with metabolic and biological processes especially response to diverse biotic and abiotic stimuli using *in silico* approach. Computational analysis based on cluster bean ESTs led to the identification of 57 miRNAs along with their targets. To the best of our knowledge, this is the first report on identification of miRNAs and their targets using ESTs in cluster bean. The miRNA related to gum metabolism was also identified. Most abundant miRNA families predicted in our study were miR156, miR172, and miR2606. The length of most of the mature miRNAs was found to be 21nt long and the range of minimal folding energy (MFE) was 5.8–177.3 (−kcal/mol) with an average value of 25.4 (−kcal/mol). The identification of cluster bean miRNAs and their targets is predicted to hasten the miRNA discovery, resulting in better knowledge of the role of miRNAs in cluster bean development, physiology, and stress responses.

**Keywords:** micro RNAs (miRNAs), cluster bean (*Cyamopsis tetragonoloba* L. Taub.), miRNA identification, miRNA targets, non coding RNAs, galactomannan

**1 INTRODUCTION**

Cluster bean (*Cyamopsis tetragonoloba* L. Taub.) popularly known as guar is a drought-tolerant, large-seeded legume, mainly adapted to semi-arid parts of the world with minimal requirements. It is commonly used as a source of food and feed in the Indian subcontinent and is a good source of nutrition, as it is a rich source of dietary fiber, folic acid, and Vitamin C (Rajaprakasam et al., 2021). India accounts for 80% of the global cluster bean production making it second-largest cash crop in India (Sahu et al., 2020). In the past two decades, the demand for cluster bean production has increased and cultivation has also been initiated in non-traditional areas like the United States, Australia, Brazil, Italy, Spain, Morocco and Germany, most recently, Russia due to the discovery of a
valuable polysaccharide-galactomannan (guar gum), in the seed endosperm (Teplyakova et al., 2019; Grigoreva et al., 2021a). The gel-like structure formed by hydrated gum has a wide utility in various industries like paper, food and agriculture, cosmetics, explosives, petrochemicals, and water purification where it is used as thickener, stiffener, stabilizer, and strengthening agent (Mudgil et al., 2014; Sharma et al., 2018; Saya et al., 2021). The demand for guar gum has increased by four times in crude oil industry since 2006 (Bhatt et al., 2016). The medicinal properties of cluster bean viz. antihyperlipidemic, antihyperglycemic, and antihypercholesterolemic have made it an important constituent of pharma industry. Recently, it has been used in nanotechnology-based drug delivery and as binder in electrochemical industry (George et al., 2019; Verma and Sharma, 2021; Kaur and Santra, 2022). The global demands for cluster bean in various industries have increased several folds in the past decade making it a crop of high economic importance.

Despite such high economic importance this crop has not been explored and characterized to its full extent at genomic level. Though, recently some studies have been reported (Kaila et al., 2017; Rawal et al., 2017; Tanwar et al., 2017; Al-Qurainy et al., 2019; Chaudhury et al., 2019; Sahu et al., 2020; Rajaparakasam et al., 2021) and the availability of the whole genome sequence which identified 34,680 protein-coding genes from 550.31 Mbp of genome will further help in accelerating the fundamental and applied research in this crop (Gaikwad et al., 2020 Unpublished). When it comes to the non-coding genome, especially microRNA the research is still at its initial stages and yet to be explored to its full potential. miRNAs are single-stranded RNAs with a length of 21–25 nucleotides (nt) that are formed from stem-loop precursors and are highly conserved (Carthew and Sontheimer, 2009; O’Brien et al., 2018). These small non-coding RNAs (miRNAs) are involved in regulation of growth, development, and metabolic pathways. In Leguminosae family, a wide range of miRNAs has been reported from nine species (Kozomara and Griffiths-Jones, 2014). miRNAs and their precursors in response to various physiological stresses have also been reported in important cereals like rice, wheat, maize, sorghum, and barley (Budak and Akpinar, 2015; Djami-Tchatchou et al., 2017; Singroha et al., 2021). All these studies showed that miRNAs play an important role in regulation of physiological responses, developmental and metabolic pathways, and evolution in plants.

When genome sequence is unavailable, EST (expressed sequence tags) and GSS (genome survey sequences) data can be utilized for identification of miRNA and their targets. In *silico* miRNA prediction and characterization are developing as faster and more efficient methods than laboratory-based cloning methods, as they are cheap and quick (Gupta et al., 2017). Such *in silico* methods based on EST and GSS have been used to identify miRNAs and their target in various crop plants including cotton, soybean, lentil, pepper, potato, sweet potato, tomato, tobacco, and horse gram (Qiu et al., 2007; Zhang et al., 2008; Kim et al., 2011; Budheswar et al., 2013; Dehury et al., 2013; Vivek, 2018; Yasin et al., 2020). However, no miRNAs for the cluster bean have been reported in miRBase, and these have yet to be properly studied. Though, few recent studies report the identification of various miRNA families in cluster bean (Tyagi et al., 2018; Sahu et al., 2018). In the present study, we used EST-based identification and characterization of conserved miRNAs belonging to different families and their probable target genes. Predicted miRNAs were also functionally annotated using network analysis and gene ontology. Understanding the role of miRNAs in the control of various biological processes necessitates the study of interactions between individual miRNAs and their target genes (Bansal et al., 2017). Further, the discovery of cluster bean miRNAs and their targets is predicted to hasten miRNA discovery, resulting in a better knowledge of the role of miRNAs in cluster bean development, physiology, and stress responses.

## 2 MATERIALS AND METHODS

### 2.1 Sequence Collection and Software Information

A total of 16,503 EST sequences of cluster bean were downloaded from the GenBank nucleotide database (http://www.ncbi.nlm.nih.gov/nuccst) available at NCBI. To identify conserved miRNAs genes and their targets, C-mii (version 1.11) programme was used (Nunnmark et al., 2012). To determine putative miRNA sequences, 4,256 non-redundant mature plant miRNA sequences of viridiplantae (green plants) were also downloaded from miRBASE. BLASTN with an e-value cut-off equal to 10 was performed against the EST sequence of cluster bean and mature miRNA sequences from miRBASE. Next, to remove protein-coding sequences BLASTX with e value b = 1e−5 was done against UniprotKB/Swissprot (release 2010-12) and UniProtKB/TrEMBL (release 2011-01) protein databases. The primary and precursor miRNAs structures were predicted using RNA database Rfam 10. UNAFold was run at a maximum base pair distance of 3,000, maximum bulge/interior loop size 30, and single read run temperature of 37°C.

### 2.2 Prediction of the Secondary Structure of Cluster Bean miRNAs

miRNA sequences that matched at least 18 nt and had a 3 nt mismatch with all known plant mature miRNAs were chosen for further study. The top BLAST hits were collected, and the BLASTX tool was used to execute the anticipated precursor sequences. Non-coding sequences were preserved, while protein-coding precursor sequences were removed. Rfam (10) was used to differentiate miRNA from other RNA families (rRNA, snRNA, and tRNA). UNAFold was used to anticipate secondary structure using parameters such as maximum base pair distance (3,000), maximum bulge/interior loop size (30), and single tread run temperature (37°C).

Using a homology search, the following criteria were utilized to find miRNAs: 1) The predicted mature miRNAs should be 19–25 nucleotides in length, 2) Maximum of four mismatches against the reference miRNA were allowed, 3) The mature miRNA should be localized within a stem-loop structure with one arm, 4) No more than five mismatches were allowed between
miRNA sequences and the guide miRNA sequence, 5) The A + U content should be high, and the secondary structure’s minimal folding free energy (MFE) and MFE index (MFEI) values should be significantly negative.

The MFEI was calculated using the following equation:

\[
\text{MFEI} = \left( \frac{\text{MFE}}{\text{length of the RNA sequence}} \right) \times 100 / (G + C)\%
\]

2.3 Prediction of miRNA Targets
C-mii and the Plant Target Prediction Tool, both available on the UEA sRNA toolkit (srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi), were used to estimate the likely target locations of detected miRNAs. miRNA-targeted miRNAs have perfectly or almost perfectly complementary sites with miRNAs and miRNAs suppress gene expression by binding to these targeted miRNAs at their complementary sites for direct mRNA cleavage or protein translation suppression (Bartel, 2004; Chen et al., 2004). This suggests that using homology search to predict miRNA targets in plants is an effective method. Target identification was carried out by comparing identified miRNA candidates to the same transcript used for miRNA identification. To predict miRNA-target genes, the following parameters were used: 1) No more than four mismatches were allowed between projected miRNAs and target genes, 2) No mismatches were allowed for the 10th and 11th positions of the complementary site since they are regarded as cleavage sites, and 3) Maximum of four GU pairs were allowed in the complementary alignment. All the steps involved in identification of miRNA and their targets in cluster bean are represented in Figure 1.

2.4 Phylogenetic Analysis of Identified miRNAs
The precursor sequences of the identified and well-known cluster bean miRNAs were aligned and phylogenetically analyzed to examine their evolutionary relationships (www.clustal.org/). The evolutionary distances were calculated using the neighbor-joining (NJ) method after 1,000 bootstrapped iterations.

2.5 Analysis of Gene Ontology for Predicted Targets
GO analysis of the identified target transcripts and pathway analysis was performed using the target annotation module in C-mii, which calculates and visualizes the distribution of selected targets, to better understand the role of cluster bean miRNAs and their regulating targets.

2.6 The miRNA-Target Gene Regulatory Network
To study co-regulated targets by miRNA families and to select miRNA-targets based on MFE value, a biological network was created between identified miRNAs and their targets.
GENEMANIA (http://genemania.org/) was used to visualize the biological network of miRNA and its target.

2.7 Identification of Transcription Factors From Expressed Sequence Tags of Cluster Bean
Homology based search against the plant transcription factor database (PlnTFDB) was performed to identify genes encoding transcription factors in the EST of cluster bean (Pérez-Rodríguez, 2010).

3 RESULTS
3.1 Identification of miRNA Using Expressed Sequence Tags Sequences
To identify and characterize the conserved miRNA in cluster bean, a thorough EST-based approach using C-mii (version1.11, Nummack et al., 2012) software was employed as genomic sequence information is limited. The software includes four steps such as sequence loading, homolog search, primary miRNA folding, and precursor miRNA folding. A total of 16,503 EST sequences of cluster bean were downloaded from the public database available at (http://www.ncbi.nlm.nih.gov/nucest) and compared to the non-redundant viridiplantae mature plant miRNA sequences (4,256) retrieved from miRBase (Supplementary Figure S1). Sequences longer than 3000 nt or containing characters other than A, T, C, G, U, and N were excluded. The homolog search module scans different plant mature miRNAs from miRBase with 16,503 EST sequences with default parameter setting. Using the plus strand, only 361 miRNA families and 670 members were identified. These sequences were selected as input for the primary miRNA folding module. Next, the primary miRNA folding was done with default parameter settings except for the BLAST e-value ≤ 1E-5 which resulted in successfully folding 3,068 sequences. These sequences were again selected for the next step i.e., the precursor miRNA folding module. A different number of sequences remained at each step (Supplementary Table S1). The precursor miRNA folding at default parameter settings except for the BLASTX e-value ≤ 1E-5 resulted in 74 miRNA candidates, belonging to 57 miRNA families predicted from cluster bean EST sequences available at NCBI. For miRNA families which were predicted from EST database were matched against cDNA library of seed development stage and out of 57 EST IDs in which different miRNA families were present, only 19 EST IDs were found to be involved in various functions (Supplementary Table S2).

3.2 Characteristics of Newly Identified Cluster Bean miRNA
To validate newly identified miRNAs, various calculated parameters were analyzed for miRNA family member, precursor length of miRNA, A + U content, G + C content, their ratios, the minimal folding free energy (MFE), and minimal folding free energy index (MFEI) for each miRNA precursor.

3.2.1 miRNA Family Members
Most of the miRNA families had one member but some of the miRNAs families were observed to have more than one number. These families were miR156 (9), miR 2606 (7), and miR172 (Supplementary Figure S2).

3.2.2 Precursor miRNA Length
The length of precursor miRNAs varied from 35 to 539 nt in length with an average value of 100 nt. miR1132 was found to have the lowest miRNA pre-miRNA length (35 nt) and miR1439 had highest miRNA length (539 nt; Supplementary Figure S3).

3.2.3 Mature miRNA Length
The length of most of the mature miRNAs was found to be 21 nt. Based on miRNA length, the predicted miRNA was divided into miRNA families which were predicted from EST database were matched against cDNA library of seed development stage and out of 57 EST IDs in which different miRNA families were present, only 19 EST IDs were found to be involved in various functions (Supplementary Table S2).

miRNA families which were predicted from EST database were matched against cDNA library of seed development stage and out of 57 EST IDs in which different miRNA families were present, only 19 EST IDs were found to be involved in various functions (Supplementary Table S2).
three groups. These groups had 19, 20, or 21 nucleotide length in mature miRNA except in two groups. First group was 19 nt long, while second group was 20 nt long, and third group was 21 nt long (Supplementary Table S3).

### 3.2.4 Minimal Folding Free Energy

The minimal folding free energy (MFE) has been considered one of the significant features for identification of miRNA. It is an important parameter for determining the secondary structure of miRNA.
pre-miRNA. Highly negative MFE indicates thermodynamically stable secondary structure of the corresponding sequences. In the present study, the range of MFE (ΔG kcal/mol) was 5.8–177.3 (−kcal/mol) with an average value of 25.4 (−kcal/mol) (Supplementary Figure S4).

3.2.5 Minimal Folding Free Energy Index
The index that differentiates pre-miRNA from other coding and non-coding RNAs and RNA fragments is MFEI. The MFEI values observed in the present study ranged from 0.59 to 1.36 −kcal/mol. The mean MFEI values were found to be 0.70 −kcal/mol (Supplementary Figure S5). The MFEI values of tRNAs, rRNAs, and mRNA are 0.64, 0.59, and 0.62–0.66 −kcal/mol, respectively which are significantly lower than the cluster bean miRNAs reported in this study, signifying that the identified cluster bean miRNAs are true miRNAs.

3.2.6 AU Content
The analysis also showed that AU contents in all identified miRNA precursors ranged from 41.09 to 83.12% with an average of 61.18% (Supplementary Figure S6).

3.2.7 GC Content
GC content was also analyzed for different miRNA families predicted in cluster beans. The GC content ranged from 18.36 to 60.45% with an average of 38.81% (Supplementary Figure S7).

3.2.8 Nucleotide Content
The nucleotide content was not uniform throughout the predicted miRNA families (Supplementary Figure S8). The A/U and C/G ratio was also calculated as given in Supplementary Table S4. The highest A/U was observed for miR1527 (1.6) and the lowest was for miRNA 444 (0.3) whereas the highest C/G ratio was for miR1527 (1.4) and the lowest for miR537 (0.2). All parameters of different characteristics of predicted miRNA families are summarized in Table 1.

3.3 Sequence Alignment and Phylogenetic Analysis of New miRNA
Primary and mature plant miRNAs are highly conserved among distantly related plant species. This high level of conservation between taxa was used as an ancillary criterion for miRNA annotation. Comparison of the precursor sequences of the predicted miRNAs with each other showed that most members could have a high degree of sequence similarity with others. Three different clusters were formed which I cluster was the largest and was divided into two sub-clusters. Sub-cluster I consists of miR5542, miR399, miR169, miR393, miR437, miR2606, miR444, miR1857, miR1313, miR5565, miR1109, miR1044, miR5641, miR1533, miR5267, miR1527, miR2666, miR5265, and miR867. Sub-cluster II consists of miR2919, miR397, miR4413, miR865, miR2634, miR773, miR5338, miR2275, miR168, miR4995, miR2093, miR5489, miR5079, miR3979, miR902, miR172, miR414, miR2082, miR838, miR2105, miR5015, miR781, miR396, and miR837 and comprised of largest number of miRNA families as in Figure 3. Cluster II consists of miR5021, miR1134, miR5658, miR156, miR1132, miR1439, miR2628 and miR1852. Cluster III, which is the smallest of all comprises miR1535, miR1167, and miR5662.

3.4 Target Identification
The target identification procedure using a computational method is simple and efficient, yielding possible targets that aid in the subsequent study. C-mii’s target prediction menu consists of four submenus: sequence loading, target scanning,
| miRNA | Targeted protein | Targeted EST ID | Description |
|-------|------------------|-----------------|-------------|
| miR156 | SPL4, SPL13, SBP2, SBP1, SPL5 | EG982824.1 GLL094_E10_036 | Squamosa promoter-binding-like protein |
| miR172 | 2AAE | EG979243.1 GLL062_C07_030 | Ubiquitin-conjugating enzyme E2 |
| miR393 | RIK20 | EG987334.1 GLL062_E02_044 | Serine/threonine-protein phosphatase |
| miR837 | CSH2 | EG990107.1 GLE087_E01_004 | ADP-ribosylation factor |
| miR837 | G3PA | EG975619.1 GLL029_A09_040 | Glyceroldehyde-3-phosphate dehydrogenase A |
| miR414 | AGO1 | EG985306.1 GLE042_A02_008 | Protein argonaute |
| miR414 | ALFIN | EG986332.1 GLE052_H02_001 | Protein argonaute |
| miR393 | 2AAG | EG985097.1 GLE041_H10_023 | Serine/threonine-protein phosphatase |
| miR837 | ARF | EG987334.1 GLL062_E02_044 | 50S ribosomal protein |
| miR837 | G3PA | EG975619.1 GLL029_A09_040 | Glyceroldehyde-3-phosphate dehydrogenase A |
| miR414 | AGO1 | EG985306.1 GLE042_A02_008 | Protein argonaute |
| miR414 | ALFIN | EG986332.1 GLE052_H02_001 | Protein argonaute |

(Continued on following page)
miRNA-target folding, and target annotation. Different miRNA targets are represented in Table 2. The acceptable length of an input sequence for target identification was extended to 20,000 nt. In the present study, 2,554 sequences previously reported as miRNA-specific target from 47 different plant species were uploaded and selected for target scanning. Using the plus strand, 6,164 sequences were identified as miRNA binding sites for 361 miRNA families and 670 members. All of them were selected as input for miRNA target folding at 37°C. About 3,068 sequences were successfully folded and these sequences were used as input for target annotation using Gene Ontology (GO). The target annotation module predicts the function for potential targets. Target identification results are provided in Table 2 with their targeted EST ID. A total of 57 predicted miRNAs from 37 miRNA families were found to regulate 623 target transcripts. miRNAs can potentially regulate many distinct genes, implying that these genes belong to several gene families engaged in various biological, cellular, and molecular processes (Dehury et al., 2013). However, several miRNAs with unclear roles were discovered (miR1104, miR1167, miR1313, miR1852, miR2866, miR2919, miR3522, miR4413, miR5267, miR5338, miR537, miR5489, miR779, miR781, miR867, miR863, miR444, miR5265, and miR5542).

miRNA414, miR5658, and miR5021 have the highest number of targets. Some miRNAs such as miR156, miR172, miR414, miR1533, and miR5021 had more than one target while some miRNA families have common targets as represented in Table 3.
To gain a better understanding of the functional characteristics of these miRNA targets, the GO annotations of all projected targets were examined. All targets were divided into three GO categories based on the ontological meanings of the GO words, as shown in Figures 4A–C: molecular function, biological process, and cellular component. Predicted targets in the molecular function category were related to nutrition reservoir activity (32%), followed by nucleotide binding and protein serine/threonine kinase activity. The bulk of the targets in the biological process category was potentially implicated in cellular activities and metabolic processes involving salt stress (18%) and proteolysis (10%). Other genes were found to be involved in a variety of signaling pathways, including the gibberellic acid signaling pathway, auxin-mediated signaling, and abscisic acid stimulation. Target genes in the cellular component category were assigned to the subcategories cytoplasm (24%), chloroplast (15%), nucleus (14%), mitochondria, and membrane. Major GO IDs were selected based on their function and classified into biotic and abiotic, and carbohydrate metabolism-related categories as represented in Supplementary Table S5.

The Gene Ontology provides the logical structure of the biological functions and their relationships to one another, manifested as a directed acyclic graph.

1) GO: 0009615 is involved in any process that results from a stimulus of the virus. This GO ID was found to be
associated with other GO IDs such as GO:0008150, GO:0050896, GO:0009605, GO:0009606, GO:0043207, GO:0051704, GO:51707.

2) GO: 0009414 relates to a change in the activity of cells as a result of water deprivation. This GO ID was found to be interconnected with GO:0008150, GO:0001101, GO:0005984, GO:0044262, GO:0009311, GO:00009628.

3) GO:0005985 relates to all the chemical reactions and pathways involving sucrose, the disaccharide fructofuranosyl-glucopyranoside. This GO ID was associated with GO:0008150, GO:0005984, GO:0044262, GO:0009311, GO:0010035, GO:0006950, GO:0009415, GO:01901700, and GO:00009628.
0005975, GO:0044238, GO:0044237, GO:0009987, GO: 0008152, GO:0071704, and GO:008152.

3.5 Gene Regulatory Networks of Cluster Bean Target Genes in GENEMANIA

Target genes of predicted miRNA families of cluster bean were used to make a regulatory network in GENEMANIA and to predict the function of these genes (Figure 5). Following genes were found and a regulatory network was predicted against *Arabidopsis thaliana*. The genes were CML 13, MIOX1, SCL1, RD21A, AGO1, GUX4, INVA, TIM13, ADH1, LACS9, CYP40, MIOX2, AGO4, CKS1, AR17, CML35, PMM1, ASD1, ADF2, PUB30, CUL1, MSRB3, DXS, AG, TIC32, LRE, PCKA, SNX1, ASD2, CKX3, YKT61, STO, EBP17 and RP2. The network represents the involvement of various additional genes and their function (Supplementary Table S6) using STRING.

Various genes comprised of different networks such as co-expression (55.01%), shared protein domain (35.32%), predicted (6.35%), and co-localization (3.32%). The given expression network indicates the involvement of various additional genes and their function (Supplementary Table S6; Ouellet et al., 2019). This led to the identification of many transcription factors, the bulk of which belonged to the MIKC MADS family protein basic leucine-zipper and the C3H family protein (Figure 6). A model is well represented in Figure 7, showing the interaction between miRNA and transcription factor genes, as well as their impact on many developmental processes and metabolic activities in plants. miRNAs and TFs predicted from the EST database enable combinatorial gene regulation with varied roles that can be used to enhance crops. The combination of miRNA and their targets, which are mostly transcription factors, resulted in the construction of a model suggesting a clear-cut link that leads to the regulation of many developmental and metabolic processes. MiR156, for example, interacts with AP2 and SPL to govern floral growth as well as the juvenile-to-adult transition. Similarly, miR5021 and miR5658 target MADS BOX, which is known to play a crucial role in leaf and flower growth. Furthermore, miR5658 targets heat stress transcription factors and is involved in drought stress and developmental processes such as growth and reproduction. miR838 targets NAC, which is essential for root growth.

4 DISCUSSION

In recent times numerous studies carried out by various researchers have demonstrated that plant miRNAs are involved in various biological and developmental processes and physiological responses. Gaining insights into miRNAs and their targets help us to understand the array of biological events underlying miRNA-mediated regulation and will further elucidate the functional importance of miRNAs (Pani and Mahapatra, 2013; Chaudhary, 2019). A substantial number of miRNAs are conserved among the plant species and it has been observed plant miRNAs and their target sites bind to each other with perfect or near to perfect complementarity (Wang et al., 2004; Schwab et al., 2005; Patanu et al., 2017). This provides an opportunity for identification of miRNAs and their targets in different crop plants by comparing/aligning with the known miRNA sequences. This approach has been widely used to identify miRNA and their targets in various crop plants (Sunkar and Jagadeeswaran, 2008; Dehury et al., 2013; Gasparis et al., 2016; Vivek, 2018). Here, we report the identification of miRNAs and their potential targets in cluster bean by mining EST database.

Cluster bean is an emerging economically important legume with wide range of industrial applications (Rajaprakasam et al., 2021). Despite, high industrial importance of the guar gum this crop is considered as orphan as little progress has been made and availability of genetic and genomic resources is very limited in this crop (Ravelombola et al., 2021). Recently, advanced molecular and genomic approaches have been utilized in this orphan crop to combat the yield losses and elucidate the function of galactomannan related genes (Al-Qurainy et al., 2019; Kumar et al., 2020; Rajaprakasam et al., 2021; Grigoreva et al., 2021a; Grigoreva et al., 2021b; Acharya et al., 2022). Attempts have also been made to sequence cluster bean genome (Grigoreva et al., 2019); however, this first sequencing attempt has a gap of 50% (Grigoreva et al., 2019).

Recent studies are now gradually overcoming the limitations, however, still limited information is available on miRNA-mediated post-transcriptional regulation of metabolic pathways in cluster bean. In few reports, deep sequencing has been applied to identify miRNAs in cluster bean (Sahu et al., 2018; Tyagi et al., 2018). Deep sequencing is the most commonly employed technique for miRNA identification in plant species whose whole genome sequence is not known (Pelaez et al., 2012; Tyagi et al., 2018; Tiwari et al., 2021). However, this technique generates huge amount of data and needs expertise for data analysis. Moreover, next-generation sequencing (NGS) is costly and not possible for routine applications (Gasparis et al., 2017).
| Input sequence Id | Precursor start/Stop | Known mature miRNA sequence | Predicted miRNA sequence | Predicted miRNA family | Homolog miRNA |
|-------------------|----------------------|-----------------------------|--------------------------|------------------------|--------------|
| EG990325.1        | 554–644              | 5′:UGACAGAAGAAAGAGAGC       | 5′:UGAGAGAAGAAAGAGAGG    | miR156                 | ath-miR156h  |
| GLE099_F01_003    | 382–429              | AC3′                      | AA3′                     | miR168                 | sof-miR168b  |
| EG985292.1        | 201–258              | 5′:UGCCCUYGGGCAAGAAGGC     | GC3′                     | miR169                 | vvi-miR169l  |
| GLL073_G12_042    | 324–379              | 5′:GAGCAAGAAGAAGACUGG      | UGU3′                    | miR172                 | gma-miR172b  |
| EG984942.1        | 369–420              | 5′:UGCAAAAGGAAGAUGAUGU     | UC3′                     | miR939                 | gma-miR939   |
| GLE040_D12_045    | 205–253              | UG3′                      | AC3′                     | miR939                 | gma-miR939   |
| EG985097.1        | 413–460              | 5′:UGCCGAAGAAGAAGACUGG     | UGU3′                    | miR939                 | pab-miR937   |
| GLE041_H10_033    | 207–248              | 5′:UGCAAAAGGAAGAUGU       | CUG3′                    | miR939                 | tae-miR939   |
| EG987334.1        | 317–428              | 5′:UCCUAAGAAGAAGAAGAUG     | UCC3′                    | miR141                 | osa-miR141   |
| GLE062_E02_004    | 137–199              | 5′:AAAGUUGAAGAAGAUGAUGU   | UUC3′                    | miR437                 | osa-miR437   |
| EG987853.1        | 311–357              | 5′:UCAGAUGCGUGCCUCAAG     | CUU3′                    | miR444                 | osa-miR444a2 |
| GLL088_H10_033    | 223–300              | 5′:UGCAAGAAGAAGAAGAUGG    | UCU3′                    | miR537                 | ptp-miR537a  |
| EG975619.1        | 632–672              | 5′:UUUGCCUCCAGUUUGU       | UGC3′                    | miR773                 | ath-miR773   |
| GLE094_G01_002    | 388–541              | 5′:UUGUGCCUAAUUUUGU        | UU3′                     | miR1044                | ppt-miR1044  |
| EG985485.1        | 380–512              | 5′:UUGUGGGAAGAUGUUGU       | AAU3′                    | miR1109                | smo-miR1109  |
| GLE045_G03_010    | 335–369              | 5′:CAUAUUGGAAAGGAGG         | GCG3′                    | miR1132                | tae-miR1132  |
| EG978745.1        | 36–91                | 5′:CAACAAACAAAGAAGAAGAGAA | GAU3′                    | miR1134                | tae-miR1134  |
| GLE05_A06_024     | 44–89                | 5′:GGGGUGUGAGAAGAAGAAGAAGA | AC3′                     | miR1167                | cre-miR1167  |
| GLE045_H05_017    | 158–251              | 5′:UACACUAUUGGAAAUUUGU      | UU3′                     | miR1313                | pta-miR1313  |
| EG899165.1        | 161–752              | 5′:UUGUGGAAGAAGGAGUGAGU    | UUU3′                    | miR1439                | osa-miR1439  |
| GLE017_B06_023    | 51–96                | 5′:UAACUCAACUCUACAA        | GGU3′                    | miR1527                | gma-miR1527  |
| EG978287.1        | 453–529              | 5′:UAAUUAUUAAUUAUAAUCAG3′ | AC3′                     | miR1533                | gma-miR1533  |
| GLL02_C09_038     | 47–191               | 5′:GUUGUUGUGGUGUGGAAGUC    | C3′                      | miR1535                | gma-miR1535  |
| GLE089_D07_029    | 96–177               | 5′:AUAUGAAGAAGAAGAAGAAGC   | C3′                      | miR1852                | osa-miR1852  |
| EG899797.1        | 116–160              | 5′:UGUGUUGAAGAAGAAGAAGAAGC | GGA3′                    | miR1857                | osa-miR1857  |
| GLE079_A10_040    | 644–700              | 5′:UGUGAAGAAGAAGAAGAAGAAGC | AGG3′                    | miR2082                | ppt-miR2082  |
| GLE811_010        | 192–278              | 5′:UCUGUGUAGAAGAAGAAGAAGC | UGU3′                    | miR2098                | osa-miR2098  |
| EG981411.1        | 393–497              | 5′:UGUGUUGAAGAAGAAGAAGAAGC | UU3′                     | miR2105                | osa-miR2105  |
| GLE010_F06_019    | 253–317              | 5′:AGAUAGUAGAAGAAGAAGAAGC | AA3′                     | miR2275                | zma-miR2275c |
| GLL01_H03_009     | 573–612              | 5′:UCAAAUUGUAGAAGAAGAAGAAGC| ACC3′                    | miR2606                | mtr-miR2606a |
| EG975952.1        | 210–282              | 5′:UAAUUGAAGAAGAAGAAGAAGAAGC| AA3′                     | miR2628                | mtr-miR2628  |

(Continued on following page)
This limitation can be overcome by using in silico approaches that have emerged as fast and cheap alternative and uses databases like EST and GSS (Vivek, 2018).

Similar strategy was applied in this study and EST database was used to mine miRNAs using computational approach. A large number of miRNA families were identified in cluster bean (Table 4) belonging to drought, cold, biotic stress tolerance, and salt-specific stress conditions. As previously reported in other species, the majority of the cluster bean precursor miRNA sequences varied in length from 35 nt to 592 nt with an average value of 100 (Wang et al., 2012; Singh et al., 2016). The length of the mature miRNAs was found to be 20–22 nt long. Similar length of mature miRNAs was reported in earlier miRNA identification studies in soybean, and switchgrass (Xie et al., 2010; Li et al., 2015). The nucleotide composition showed that at the first position at 3′ end of mature miRNA uracil was pre-dominant, similar to observations made by Zhang et al. (2006). It was observed that the average AU content (61.18%) was higher than the GC content (38.18). The higher AU content depicts a comparatively less stable pre-miRNA secondary structure that is easily recognized by RISC complex and converted into mature miRNA. The minimal folding free energy (MFE) is an important determining factor for the stability of the secondary structure. The lower the values of MFE, the higher the thermodynamic stability of the secondary structure (Prabu and Mandal, 2010). In the present study, the MFE value was found to be 25.4 (−kcal/mol). The MFE reported in this study was significantly lower than MFE value of earlier reported miRNAs in cluster bean, where the most stable miRNA, cte_miR1134 has MFE of −72.6 kcal/mol (Sahu et al., 2018). Similarly, the MFE of miRNA families reported by Tyagi et al., 2018 in different tissue-specific miRNAs of cluster bean was quite higher. The MFE values of miRNAs reported in other legumes like peanut (−50.01 kcal/mol; Chi et al., 2011), French bean (−35.0 to −51.2 kcal/mol; Peláez et al., 2012), chickpea (−50.1419 kcal/mol; Kohli et al., 2014), and lentil (−44.07 kcal/mol; Hosseini et al., 2021). Nowadays, to differentiate miRNAs from other RNAs an MFEI value of higher than 0.85 is a potential criterion (Wang et al., 2012). The cluster bean pre-miRNAs identified in the present

### Table 4 | (Continued) List of predicted miRNA families in cluster bean.

| Input sequence Id | Precursor start/Stop | Known mature miRNA sequence | Predicted miRNA sequence | Predicted miRNA family | Homolog miRNA |
|--------------------|-----------------------|----------------------------|--------------------------|------------------------|----------------|
| EQ982820.1         | 5′-UUUACUUCCGAAUUUGUG   | 5′-UUUACUUCCGAAUUUGUG      | miR2834                  | mtr-miR2834            |
| GLE027_H07_025     | 5′-UCGAGUGGUGGUGGUGGUGG | AAC-3′                     | miR1134                  | mtr-miR1134            |
| GLL09_G09_024      | 5′-AAAGGAGGAGGAGGAGGAGG | UAC-3′                     | miR995                   | gma-miR995            |
| EG984029.1         | 5′-UCAGCGUCGCGCUGCGCGC | UAU-3′                     | miR5015                  | ath-miR5015            |
| GLE016_B04_015     | 5′-UAAGGUAGGAGGAGGAGGAG | AUU-3′                     | miR5015                  | ath-miR5015            |
| GLE094_B08_026     | 5′-AAAGGAGGAGGAGGAGGAGG | UAU-3′                     | miR5015                  | ath-miR5015            |
| EG982778.1         | 5′-UCCGUGGUGGUGGUGGUGG  | UAU-3′                     | miR5015                  | ath-miR5015            |
| GLE031_E11_044     | 5′-AGAGGUGACGCAUGUGAUGA | AG-3′                      | miR5015                  | ath-miR5015            |
investigation showed MFEI values ranging from 0.59 to 1.39 (±kcal/mol) with an average value of 0.70 (±kcal/mol). The MFEI values reported in this study were similar to that of wheat (0.79–1.85±kcal/mol; Gasparris et al., 2017) and lentil (±1.02 kcal/mol; Hosseini et al., 2021).

Cluster bean miRNAs were found to be involved in plant development, defense mechanisms, and signal transduction. The projected miRNAs’ functional significance can be easily appreciated by obtaining more information about their target genes. Serine/threonine-protein kinase and ubiquitin carrier protein were anticipated to be the targets of miR172, miR5021, miR156, and miR1533, respectively. Previous reports indicate that ubiquitin carrier proteins plays a significant role in the regulation of various physiological processes such as the recycling of aberrant proteins, metabolic regulation, cell cycle control, and transcription factor (TF) activation (Dreher and Callis, 2007). Serine/threonine-protein kinase plays an important function in signal transduction pathways that contributes to plant defense under both biotic and abiotic stresses (Pais et al., 2009). It has been reported that miRNAs target TFs, signal transduction factors and metabolic transporters (You et al., 2020).

Cluster bean network analysis of projected miRNAs and their target genes revealed information about the co-regulation of numerous target genes. TFs and miRNAs are both thought to be important regulators of transcriptional events and ultimately, the plant metabolic process (Jangra et al., 2018). TFs regulate gene transcription in the promoter area of the gene, whereas miRNAs regulate gene post transcription in the 3’-untranslated region of the gene, and their respective targets form interconnected functional networks that are critical for the execution of any metabolic activity. Annotation of miRNA-targeted unigenes revealed enrichment of a total 39 TFs such as bHLH (250), Myb (150), Myb related (126), ERF (130), bZip and C2H2 (107), WRKY (100) and NAC (99) (Tyagi et al., 2018). A total of 35 novel TFs were identified which belonged to MIKC_MADS family protein, basic leucine-zipper, and C3H family protein, NAC and WRKY classes. Plant MADS-box genes have been identified as regulators of floral organ identity to control various developmental processes such as the determination of meristem identity of vegetative, inflorescence and floral meristems, root growth, ovule and female gametophyte development, flowering time, fruit ripening and dehiscence (Masiero et al., 2011; Castelán-Muñoz et al., 2019). Certain miRNA-targeted TFs such as NAC and ZF-TFs, implicate their involvement in cellulose synthase-like proteins which is a key factor controlling galactomannan synthesis. The presence of such TFs among the identified miRNA further proves their role in galactomannan pathway. The present study demonstrates the information of miR 5658 family targeting CSL G1, G2, G3 proteins in cluster bean. In both the primary and secondary walls of Arabidopsis, CSL genes are responsible for the majority of glucomannan production (Goubet et al., 2009; Gigli-Bisceglia et al., 2020). In plants, miRNAs function as a negative regulator at the post-transcriptional gene level, either facilitating the cleavage of target transcripts or suppressing their translation, resulting in significant changes in metabolic pathway activity. These are known to regulate cell signaling, oxidative stress, abiotic and biotic stress response, and the development of different tissues, including leaf, stem, anther, root, and flower. Cleavage of the target transcript appears to be the most important and common method of gene control by miRNAs in plants (Catalanotto et al., 2016). The in silico identification of predicted miRNAs and their target genes from the EST database suggests a regulatory role in developmental stages. The EST database used in the present study is derived from early and late seed development-specific cDNA libraries. It is the time when galactomannan is synthesized and deposited in seed as a food reserve. Since TFs control gene expression, understanding their function will involve detailed genome-wide functional analysis. Nonetheless, the study provides preliminary information on miRNAs and TFs interplay in cluster bean. In silico data analysis gives information regarding miRNA identification and their interaction with TFs and target prediction which can be further supported by wet lab. experiment. We have identified miRNA- SSRs and observed DNA polymorphism in cluster bean germplasm (96 genotypes including commercial varieties and wild species). An amplicon of 450 bp and 120 bp specific to miR169 was observed in cluster bean varieties HG 2.20 and RGC 471, respectively. HG 2.20 is tolerant against bacterial leaf blight, alternaria blight, and root rot, while RGC 471 is tolerant to bacterial blight. The miR169 has been found to be associated with biotic stress tolerance (Song et al., 2018; Šečić et al., 2021). The appearance of amplicon specific to miR169 in these varieties clearly demonstrated their role in disease resistance in cluster bean (Unpublished data).

The miRNAs identified in this study are involved in regulating several genes and further gene expression studies are needed which will aid in the discovery of a new dimension of the miRNA regulatory network during plant growth and development in cluster bean.

5 CONCLUSION

Using ESTs from the NCBI database, 57 putative miRNA families were identified in cluster bean, with MFEI ranging from 0.59 to 1.36 (kcal/mol). A total of 623 target genes were predicted for 57 possible miRNA families, and their expression was suppressed either by cleavage or translational repression. Many of them were implicated in various metabolic processes. Furthermore, miRNA target prediction found that most of these genes are classified as transcription factors. A functional study of the targets revealed that they were engaged in the process of plant development. Most of the anticipated targets were shown to be regulated by several miRNAs, according to network analysis. This study provides an important view of conserved and novel miRNAs, their precursors and targets, and associated TFs. This will help the researchers working on cluster bean miRNAs in unraveling the complex gene regulatory molecular networks. The miRNAs identified through in silico approach need further experimental validation.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.
AUTHOR CONTRIBUTIONS

NY conceived and designed the research. VC carried out the experiments and collected the data. VC, SJ, and NY analyzed the data. VC and SJ wrote the manuscript and NY edited and improved the manuscript. All the authors have read and approved the final manuscript.

FUNDING

The research was funded by the Director of Research, CCS Haryana Agricultural University, Hisar through Research Grant No. C(a) BMB-6-Plan (Agri.).

REFERENCES

Aharya, B. R., Sandhu, D., Dueñas, C., Ferreira, J. F. S., and Grover, K. K. (2022). Deciphering Molecular Mechanisms Involved in Salinity Tolerance in Guar (Cyamopsis tetragonoloba (L.) Taub.) Using Transcriptome Analyses. *Plants* 11 (3), 291. doi:10.3390/plants11030291

Al-Qurainy, F., Alshameri, A., Gaafar, A. R., Khan, S., Nadeem, M., Alameri, A. A., et al. (2019). Comprehensive Stress-Based De Novo Transcriptome Assembly and Annotation of Guar (Cyamopsis tetragonoloba (L.) Taub.): An Important Industrial and Forage Crop. *Int. J. Genomics*. 2019, 7295859. doi:10.3390/plants11030291

Bansal, A., Singh, T. R., and Chauhan, R. S. (2017). A Novel miRNA Analysis Framework to Analyze Differential Biological Networks. *Sci. Rep.* 7 (1), 14064. doi:10.1038/s41598-017-14573-x

Bartel, D. P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* 116 (2), 281–297. doi:10.1016/S0092-8674(04)00405-5

Bhatt, R. K., Sharma, R., Malha, H. R., and Kalia, R. K. (2016). Guar Production, Utilization and Marketing: Current Scenario and Future Prospects. Jodhpur: Central Arid Zone Research Institute, 1–40.

Budak, H., and Akpinar, B. A. (2015). Plant miRNAs: Biogenesis, Organization and Origins. *Front. Integr. Genomics*. 15 (5), 523–531. doi:10.1016/j.fijm.2015-0451-2

Budhwar, D., Debashis, P., Sahu, J., Sahu, M., Sarma, K., Barooah, M., et al. (2013). *In Silico* Identification and Characterization of Conserved miRNAs and Their Target Genes in Sweet Potato (*Ipomoea batatas*) Expressed Sequence Tags (ESTs). *Plant Signal. Behav.* 8 (12), e26543. doi:10.4161/psb.26543

Carthey, R. W., and Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136 (4), 642–655. doi:10.1016/j.cell.2009.01.035

Castelán-Muñoz, N., Herrera, J., Cajero-Sánchez, W., Arrizubieta, M., Trejo, C., García-Ponce, B., et al. (2019). MADS-box Genes Are Key Components of Genetic Regulatory Networks Involved in Abiotic Stress and Plastic Developmental Responses in Plants. *Front. Plant Sci.* 10, 853. doi:10.3389/fpls.2019.00853

Catalanotto, C., Cogoni, C., and Zardo, G. (2016). MicroRNA in Control of Gene Expression: an Overview of Nuclear Functions. *Int. J. Mol. Sci.* 17 (10), 1712. doi:10.3390/ijms17101712

Chaudhary, V. (2019). *Development of Micro-RNA and EST-SSR Markers for Diversity Analysis in Cluster Bean (Cyamopsis tetragonoloba (L.) Taub.)* Ph.D. Thesis. Hisar: CCS Haryana Agricultural University.

Chaudhary, A., Kalia, T., and Gaikwad, K. (2019). Elucidation of Galactomannan Biosynthesis Pathway Genes through Transcriptome Sequencing of Seeds Collected at Different Developmental Stages of Commercially Important Indian Varieties of Cluster Bean (*Cyamopsis Tetragonoloba*). *Sci. Rep.* 9 (1), 11539. doi:10.1038/s41598-019-48072-w

Chen, C.-Z., Li, L., Lodish, H. F., and Bartel, D. P. (2004). MicroRNAs Modulate Hematopoietic Lineage Differentiation. *Science* 303 (5654), 83–86. doi:10.1126/science.1091903

ACKNOWLEDGMENTS

The authors thank the Director of Research and Head of the Department of Molecular Biology, Biotechnology and Bioinformatics, CCS Haryana Agricultural University, Hisar for providing necessary support and facilities for this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.930113/full#supplementary-material

Chen, X., Yang, Q., Chen, X., Wang, J., Pan, L., Chen, M., et al. (2011). Identification and Characterization of microRNAs from Peanut (*Arachis hypogaea* L.) by High-Throughput Sequencing. *Flora* 206 (11), e27530. doi:10.1016/j.flora.2011.02.023

Chen, X., Yang, Q., Chen, X., Wang, J., Pan, L., Chen, M., et al. (2011). Identification and Characterization of microRNAs and Their Target Genes in Sweet Potato (*Ipomoea batatas*) Expressed Sequence Tags (ESTs). *Plant Signal. Behav.* 8 (12), e26543. doi:10.4161/psb.26543

Chi, X., Yang, Q., Chen, X., Wang, J., Pan, L., Chen, M., et al. (2011). Identification and Characterization of microRNAs from Peanut (*Arachis hypogaea* L.) by High-Throughput Sequencing. *Flora* 206 (11), e27530. doi:10.1016/j.flora.2011.02.023

Catalanotto, C., Cogoni, C., and Zardo, G. (2016). MicroRNA in Control of Gene Expression: an Overview of Nuclear Functions. *Int. J. Mol. Sci.* 17 (10), 1712. doi:10.3390/ijms17101712

Chaudhary, V. (2019). *Development of Micro-RNA and EST-SSR Markers for Diversity Analysis in Cluster Bean (Cyamopsis tetragonoloba (L.) Taub.)* Ph.D. Thesis. Hisar: CCS Haryana Agricultural University.

Chaudhary, A., Kalia, T., and Gaikwad, K. (2019). Elucidation of Galactomannan Biosynthesis Pathway Genes through Transcriptome Sequencing of Seeds Collected at Different Developmental Stages of Commercially Important Indian Varieties of Cluster Bean (*Cyamopsis Tetragonoloba*). *Sci. Rep.* 9 (1), 11539. doi:10.1038/s41598-019-48072-w

Chen, C.-Z., Li, L., Lodish, H. F., and Bartel, D. P. (2004). MicroRNAs Modulate Hematopoietic Lineage Differentiation. *Science* 303 (5654), 83–86. doi:10.1126/science.1091903
Hosseini, S. Z., Ismaili, A., Nazarian Firoozabad, F., Fallahi, H., and Rezaeinejad, A. (2021). Identification of Heat Stress-Responsive microRNAs in Lentil (Lens culinaris) and Their Target Genes. *Iran. J. Biol. Sci.* 34 (1), 143–154.

Jangra, S., Chaudhary, V., and Yadav, N. R. (2018). “Transcription Factors and microRNA Interplay: A New Strategy for Crop Improvement,” in *Transcriptional and Post-transcriptional Regulation*. Editor K. Ghedira (London: Intech Open). doi:10.5772/intechopen.75942

Kalia, T., Chaduvula, P., Rawal, H., Saxena, S., Tyagi, A., Mithra, S., et al. (2017). Chloroplast Genome Sequence of Clusterbean (*Cyamopsis tetragonoloba* L.): Genome Structure and Comparative Analysis. *Genes* 8 (9), 212. doi:10.3390/genes8090212

Kaur, S., and Santra, S. (2022). Application of Guar Gum and its Derivatives as Green Binder/Seperator for Advanced Lithium-Ion Batteries. *ChemistryOpen* 11 (2), e202100209. doi:10.1002/open.202100209

Kim, H.-J., Baek, K.-H., Lee, B.-W., Choi, D., and Hur, C.-G. (2011). *In Silico* identification and Characterization of microRNAs and Their Putative Target Genes in Solanaceae Plants. *Genome* 54 (2), 91–98. doi:10.1119/G10-104

Kohli, D., Joshi, G., Deokar, A. A., Bhardwaj, A. R., Agarwal, M., Katiyar-Agarwal, S., et al. (2014). Identification and Characterization of Wilt and Salt Stress-Responsive microRNAs in Chickpea through High-Throughput Sequencing. *PLoS One* 9 (10), e108851. doi:10.1371/journal.pone.0108851

Konstantinov, A., and Griffiths-Jones, S. (2014). miRBase: Annotating High Confidence microRNAs Using Deep Sequencing Data. *Nucl. Acids Res.* 42 (D1), D68–D73. doi:10.1093/nar/gkt1181

Kumar, S., Palve, A. S., Patel, S. K., Selvanayagam, S., Sharma, R., and Rathore, A. (2020). Development of Genomic Microsatellite Markers in Cluster Bean using Next-Generation DNA Sequencing and Their Utility in Diversity Analysis. *Carr. Plant Biol.* 21, 100134. doi:10.1186/s13197-020-00216-x

Li, W., Wang, P., Li, Y., Zhang, K., Ding, F., Nie, T., et al. (2015). Identification of MicroRNAs in Response to Different Day Lengths in Soybean Using High-Throughput Sequencing and qRT-PCR. *PLoS ONE* 10 (7), e0132621. doi:10.1371/journal.pone.0132621

Masiero, S., Colombo, L., Grini, P. E., Schnittger, A., and Kater, M. M. (2011). The Emerging Importance of Type I MADS Box Transcription Factors for Plant Reproduction. *Plant Cell* 23 (3), 865–872. doi:10.1105/tpc.110.007137

Mudgil, D., Barak, S., and Khatkar, B. S. (2014). Guar Gum: Processing, Properties and Food Applications-A Review. *J. Food Sci. Technol.* 51 (3), 409–418. doi:10.1007/s13197-011-0522-x

Nunnark, S., Mhuantong, W., Ingsriswang, S., and Wichadakul, D. (2012). C-mii: Throughput Sequencing and qRT-PCR. *Indian J. Genet. Plant Breed.* 72 (5), 1007/s10681-012-0229-2

Rawal, H. C., Kumar, S., Mithra S V, A., Solanke, A. U., Nigam, D., Saxena, S., et al. (2017). High Quality Unigenes and Microsatellite Markers from Tissue Specific Transcriptome and Development of a Database in Clusterbean (*Cyamopsis tetragonoloba*, L. Taub). *Genes* (Basel) 8 (11), 313. doi:10.3390/genes8110313

Sahu, S., Rao, A. R., Pandey, J., Gaikwad, K., Ghoshal, S., and Mohapatra, T. (2018). Genome-wide Identification and Characterization of lncRNAs and miRNAs in Cluster Bean (*Cyamopsis tetragonoloba*). *Genome* 667 (667), 112–121. doi:10.1016/j.genetics.2018.05.027

Sahu, S., Sahu, T. K., Ghosal, S., Gaikwad, K., and Rao, A. R. (2020). Computational Analysis of SNPs and INDELs in Cluster Bean Cultivars Involving in Multiple Trait Through Identi fication and Characterization of Wilt and Salt Stress-Responsive microRNAs in Chickpea through High-Throughput Sequencing.

Qi, X., C., Xie, F. L., Zhu, Y. Y., Guo, K., Huang, S. Q., Nie, L., et al. (2007). Computational Identification of microRNAs and Their Targets in Gossypium hirsutum Expressed Sequence Tags. *Gene* 395 (1-2), 49–61. doi:10.1016/j.gene.2007.01.034

Rajaprakasam, S., Rahman, H., Karunagaran, S., Bapu, K., J.R., G., Kulandivelu, G., et al. (2021). Comparative Transcriptome and Metabolome Profiling in the Maturing Seeds of Contrasting Cluster Bean (*Cyamopsis tetragonoloba* L. Taub) Cultivars Identified Key Molecular Variations Leading to Increased Gum Accumulation. *Gene* 791, 145727. doi:10.1016/j.gene.2021.145727

Rawelombola, W., Manley, A., Adams, C., Trostle, C., Ale, S., Shi, A., et al. (2021). Genetic and Genomic Resources in Guar: A Review. *Euphytica* 217 (11). doi:10.1007/s10681-021-02929-2

Reyes, J. L., et al. (2012). Identification of microRNAs and Their Targets in *Camellia sinensis*. *Mol. Biotechnol.* 49 (2), 100134. doi:10.1186/1471-2229-8-37

Sunkar, R., and Jagadeeswaran, G. (2008). *In Silico* Identification of Conserved microRNAs in Large Number of Diverse Plant Species. *BMC Plant Biol.* 8, 37. doi:10.1186/1471-2229-8-37

Tanwar, U. K., Pruthi, V., and Randhawa, G. S. (2017). RNA-seq of Guar (*Cyamopsis tetragonoloba*, L. Taub) Leaves: De Novo Transcriptome Assembly, Functional Annotation and Development of Genomic Resources. *Front. Plant Sci.* 8, 91. doi:10.3389/fpls.2017.00091

Teplyakova, S. B., Volkov, V. A., Dzyubenko, E. S., et al. (2021). Comparative Transcriptome and Metabolome Profiling in the Maturing Seeds of Contrasting Cluster Bean (*Cyamopsis tetragonoloba* L. Taub) Cultivars Identified Key Molecular Variations Leading to Increased Gum Accumulation. *Gene* 791, 145727. doi:10.1016/j.gene.2021.145727

Tyagi, A., Nigam, D., Mithra, S. V. A., Solanke, A. U., Singh, N. K., Sharma, T. R., et al. (2018). Genome-Wide Discovery of Tissue-Specific miRNAs in Clusterbean (*Cyamopsis tetragonoloba*) Indicates Their Association With
Galactomannan Biosynthesis. *Plant Biotechnol. J.* 16 (6), 1241–1257. doi:10.1111/pbi.12866

Verma, D., and Sharma, S. K. (2021). Recent Advances in Guar Gum Based Drug Delivery Systems and Their Administrative Routes. *Int. J. Biol. Macromol.* 181, 653–671. doi:10.1016/j.ijbiomac.2020.03.087

Vivek, A. T. (2018). In Silico Identification and Characterization of microRNAs Based on EST and GSS in Orphan Legume Crop, *Lens culinaris* Medik (Lentil). *Agri Gene* 8, 45–56. doi:10.1016/j.aggene.2018.05.003

Wang, Y.-J., Reyes, J. L., Chua, N.-H., and Gaasterland, T. (2004). Prediction and Identification of *Arabidopsis thaliana* microRNAs and Their mRNA Targets. *Genome Biol.* 5 (9), R65. doi:10.1186/gb-2004-5-9-r65

Wang, J., Yang, X., Xu, H., Chi, X., Zhang, M., and Hou, X. (2012). Identification and Characterization of microRNAs and Their Target Genes in *Brassica oleracea*. *Gene* 505 (2), 300–308. doi:10.1016/j.gene.2012.06.002

Xie, F., Frazier, T. P., and Zhang, B. (2010). Identification and Characterization of microRNAs and Their Targets in the Bioenergy Plant Switchgrass (*Panicum virgatum*). *Planta* 232, 417–434. doi:10.1007/s00425-010-1182-1

Yasin, J. K., Mishra, B. K., Pillai, M. A., Verma, N., Wani, S. H., Elansary, H. O., et al. (2020). Genome Wide In-Silico miRNA and Target Network Prediction from Stress Responsive Horsegram (*Macrotyloma uniflorum*) Accessions. *Sci. Rep.* 10, 17203. doi:10.1038/s41598-020-73140-x

You, G., Zu, B., Wang, B., Fu, Q., and Li, F. (2020). Identification of miRNA-miRNA-TFs Regulatory Network and Crucial Pathways Involved in Tetralogy of Fallot. *Front. Genet.* 11, 552. doi:10.3389/fgene.2020.00552

Zhang, B., Pan, X., Cobb, G. P., and Anderson, T. A. (2006). Plant microRNA: A Small Regulatory Molecule with Big Impact. *Dev. Biol.* 289 (1), 3–16. doi:10.1016/j.ydbio.2005.10.036

Zhang, B., Pan, X., and Stellwag, E. J. (2008). Identification of Soybean microRNAs and Their Targets. *Planta* 229 (1), 161–182. doi:10.1007/s00425-008-0818-x

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Chaudhary, Jangra and Yadav. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**