Identification of miRNAs and their corresponding mRNA targets from chickpea root infected with *M. ciceri* and functional characterization of candidate miRNAs by overexpression in chickpea roots

Manish Tiwari, Baljinder Singh, Manisha Yadav, Vimal Pandey and Sabhyata Bhatia*

National Institute of Plant Genome Research, Jawaharlal Nehru University Campus, Aruna Asaf Ali Marg, New Delhi, 110067, India.

Corresponding author

***Dr. Sabhyata Bhatia National Institute of Plant Genome Research Jawaharlal Nehru University Campus Aruna Asaf Ali Marg, New Delhi 110067, India. Email: sabhyatabhatia@nipgr.ac.in**

Running title: miRNA-target pair identification and characterization in chickpea nodulation.

Keywords: chickpea, Nodule, miRNA, PARE, bacterial small-RNA, Hairy root transformation.

+ indicates the equal contribution

Abstract

Nitrogen fixation takes place in root nodules which involves bacterial colonization, organogenesis and nitrogen fixation. Investigations related to global analysis of miRNAs mediated regulation of symbiosis in crop plants is limited. To gain a deeper insight into the miRNAs regulating gene cascade during chickpea nodulation an Illumina sequencing of miRNA library from roots subjected to infection with *Mesorhizobium ciceri* was sequenced. Using stringent criteria of miRNA annotation, a set of 91 miRNAs were identified that comprised of 84 conserved, 7 novel miRNAs with 9 pairs being polycistronic. Further, eighteen legume specific and 13 chickpea specific miRNAs were also obtained that may have specific roles in symbiosis. Interestingly, phylogenetic analysis of the precursor sequences revealed clustering of distinct miRNAs representing a close ancestry. In silico analysis also established 3 different mode of biogenesis of miRNAs. Mapping of miRNA reads to bacterial genomes helped to predict bacterial smallRNAs that may be putatively regulating host genes. Further for identification of *in-vivo* targets of miRNAs, 4 degradome libraries were sequenced. Analysis revealed 245 target transcripts that were specifically cleaved during nodule stages, with a significant number being transcription factors. qRT-PCR based expression profiling in different chickpea tissues was carried out to validate the antagonistic expression of the miRNA-target pairs. For functional characterization, 4 miRNAs, miR171f, miR172c, miR394 and miR1509, were ectopically expressed in chickpea roots by hairy root transformation that resulted in significant changes in nodule numbers. Results indicated the roles of miR171f, miR394 and miR1509 in regulating novel targets *Nodulation receptor kinase, Histidine phosphotransferase* and *Adenylalte kinase* respectively being reported for the first time that may be the key regulators of chickpea nodulation. This study not only provides an overview of the miRNAs and their targets involved in chickpea-rhizobia symbiosis but also provides several leads into novel and nodule specific miRNAs and their targets for further investigation.

Introduction

The last 2 decades have witnessed the emergence of small RNA molecules which regulate the gene expression by silencing especially post transcriptional gene silencing (PTGS) (Hamilton and Baulcombe, 1999; Mette et al., 2000). One of the indispensable components of the PTGS include microRNAs (miRNAs) that are a class of small endogenous RNAs, approximately 20-24 nt in length. They can negatively regulate the expression of a gene either by DNA methylation, mRNA degradation or translational inhibition (Voinnet, 2009; Bartel, 2004). Some miRNAs also act by production of secondary small interfering RNAs (siRNA and tasiRNA) (Chen et al., 2010). RNA polymerase II transcribes microRNA encoding genes, which undergo further processing by 5' capping, splicing, and polyadenylation at 3'end to form pri-miRNA. These are further processed/cleaved to produce a precursor RNA (Pre-miRNA) with a stem-loop structure, which is the cleaved by DCL1 to give rise to the mature miRNA-miRNA* duplex having a two nucleotide overhang at 3' end (Mishra and Mukherjee, 2007). The mature miRNAs bring about PTGS through the small RNA guided cleavage that is mediated using the Argonaute protein that has an endonuclease domain and a RNA binding domain (Song et al., 2004). A RISC complex is formed which comprises of mature miRNA and AGO protein. This RISC complex along with the region of complementarity between the miRNA and target mRNA determines whether the cleavage pathway or translational repression of target mRNA has to occur (van den Berg et al., 2008). miRNA regulate various developmental processes such as cell proliferation, stress responses, nutrition, metabolism and development etc. (Zhang and Wang, 2014; Rogers and Chen, 2013; Chuck et al., 2009).

The focus of this study is to unravel mechanisms of gene regulation in chickpea nodules. Nodule development is a tightly regulated molecular dialogue between the host plant and the rhizobia. Though several mechanisms exist to govern and modulate the gene cascades during the symbiosis process, regulation mediated through miRNAs has been the focus of attention recently. Studies in model crop *Medicago* revealed that overexpression of miRNA166 and miRNA169 led to reduction of nodule formation by cleavage of their respective targets i.e. class III HD-zip transcription factor and *MtHAP2.1* (Boualem et al., 2008; Combier et al., 2006). Similarly, it has been shown that the GRAS family transcription factor *NSP2* (Nodulation Signaling Pathway 2) which is involved in nod factors signaling is targeted by miRNA171 and ectopic expression of miR171 resulted in a significant decrease in nodule numbers (Hofferek et al., 2014).

Recently, due to the emergence of next-generation sequencing technologies and the development of *in-silico* approaches, it has become possible to sequence miRNA libraries and predict conserved and novel miRNAs. Further, deep sequencing has also facilitated the development of an approach for the high-throughput identification of miRNA targets. This approach known as Parallel Analysis of RNA Ends (PARE) analyses the RNA degradome generated through miRNA derived cleavage products (German et al., 2009). Several studies reporting the sequencing of root nodule specific microRNA libraries are available in legumes such as *Medicago*, common bean and soybean (Lelandais-Brière et al., 2009; Devers et al., 2011; Formey et al., 2014, 2015; Wang et al., 2009; Subramanian et al., 2008), however very few studies of the miRNA target identification by PARE library sequencing are available from nodules. Therefore, in this study, deep sequencing of the miRNA library from chickpea nodules was done and the conserved and novel miRNAs were predicted. The miRNA targets were also identified and annotated by high-throughput sequencing of the RNA degradome libraries from different stages of nodule development. Based on the combined analysis of this data, miRNAs with nodulation specific targets were functionally characterized through ectopic expression using hairy root transformation in chickpea.

Materials and Methods

Plant Material

Root and nodule tissues were harvested from chickpea (*Cicer aritienum)* cv BGD256. Briefly seeds were surface sterilized with 0.1% $HgCl_2$ and germinated in dark at room temperature. Germinated seeds were transferred to 1% Agar plate for 4 days. After 4 days growth period they were inoculated with *Mesorhizobium ciceri* strain TAL620. *M. ciceri* was cultured in Yeast Mannitol Broth and grown at 30 °C for 3 d. Secondary culture with an OD_{600} (1.0) was used for infection. Infected tissues were collected at 1, 3, 6, 12, 24 hpi (hours post infection) and 3, 7, 14, 21 and 28 dpi (days post infection). Root at similar time points without infection served as control tissue.

RNA extraction and library construction

Total RNA was isolated from rhizobial infected roots using LiCl precipitation method (Tiwari et al., 2019). The concentration and integrity of RNA were examined using Ribogreen method followed by Agilent 4200 Tapestation system (Agilent Technologies, Santa Clara, CA). RNA with RIN values >7.0 was selected for the following analysis. Library preparation was

performed following NEBNext® Multiplex Small RNA Library Prep Set for Illumina. Briefly, 1 µg of total RNA was used for 3` adapter ligation by incubating the samples at 25 ˚C for one hour followed by hybridization with RT primer which is required to avoid the adapter-dimer formation. Next the samples were taken for 5` adapter ligation in order to capture the complete small RNA population. The samples were further preceded for RT using ProtoScript II Reverse Transcriptase followed by PCR amplification with 12 cycles. Next, the PCR amplified cDNA constructs were purified using QIAGEN mini elute. miRNA selection was performed by loading the samples on 6% TBE gels, a single band in between 140 bp to 160 bp was excised and the library from the gel was eluted by passing the sample on to Corning gel filtration columns and precipitated the eluate using 1 µl Linear Acrylamide, 25 µl 3M sodium acetate, pH 5.5 and 750 µl of 100% ethanol. The final library was validated for the presence of ~145 bp peak on Agilent 2100 Bioanalyzer system using DNA HS chip. The sequencing was done using Illumina Next Seq 500 Sequencer with 75 SE chemistry.

For degradome sequencing, total RNA isolated at different time intervals from root ranging from (1hr to 28 days) and nodule tissue were pooled and sequenced. Root tissues were pooled into one and nodule tissues were pooled for early infection from 0 hpi to 3dpi, middle and late infection (3 dpi to 28 dpi) and a cumulative pooled for infection of all the stages. PolyA RNA was purified using OligodT Dynabeads followed by 5'sRNA adapter ligation. The ligated product was fragmented and purified which was subsequently treated for phosphate removal from 3' end. After phosphate removal, the 3'sRNA adapter was ligated. Reverse transcription of the RNA library accompanied with PCR and finally sequencing was performed using Illumina GA II sequencing system.

microRNA library analysis

Reads obtained were quality filtered. Filtered reads obtained were trimmed of the adapter sequences and reads with length between 20 to 24 nt were retained and used for mapping on to the genome of chickpea, *M. truncatula,* pigeonpea, soybean, Arabidopsis, rice, *L. japonicus, Arachis hypogea, Arachis duranensis, Arachis ipaensis, Glycine soja, Phaseolus vulgaris, Trifolium pratense, Lupinus angustifolius, Vigna angularis* and *Vigna unguiculata.* The genome sequences of the above-mentioned legume species were obtained from Legume Information System (https://legumeinfo.org/). Sequences that found match with those of tRNA, rRNA and snRNAs were excluded from downstream analysis. microRNA abundance was normalized into transcripts per million (TPM) for comparison of expression pattern.

Similarly the reads were also mapped onto bacterial genomes of *Agrobacterium tumefaciens, Bradyrhizobium japonicum (USDA 6), E. coli (strain K12 substr. MG1655), Mesorhizobium ciceri (biovar biserrulae WSM1271), Mesorhizobium ciceri (CC1192), Sinorhizobium meliloti (1021)* and *Rhizobium tropici (CIAT 899).* For comparison the reads of a nodule library from common bean were also mapped on same bacterial genomes (Formey et al., 2015).

microRNA prediction and chromosomal map

In order to predict the miRNAs from chickpea nodules ShortStack and miRDeep-P with core algorithms and plant specific criteria were used (Axtell, 2013; Yang and Li, 2018). In both the algorithms reads were mapped to chickpea and other plant genomes and flanking 250 bp window was used for RNA secondary structure prediction. The folded RNA secondary structures which fulfill all the latest criteria for miRNA annotation were chosen for mature miRNA prediction (Axtell and Meyers, 2018).

To map the mature miRNAs on chickpea chromosomes the genomic locations of the precursor miRNA were used to position the miRNAs on the eight chickpea chromosomes.

Alignment, phylogenetic analysis and biogenesis of miRNAs

The precursor sequences of the miRNAs identified were aligned using MUSCLE program and a neighbor joining phylogenetic tree was constructed in MEGA using default parameters (Kumar et al., 2018).

The precursor sequences of the miRNAs predicted from the genome of other legumes using the nodule library reads were aligned using MUSCLE program and viewed in Jalview (https://www.jalview.org/). The modes of biogenesis of mature miRNA from precursor were predicted based on these alignments.

microRNA target analysis

PARE libraries of nodule (early, mid-late and pooled nodule) and root prepared as mentioned above were used for target prediction. Reads obtained after degradome sequencing were filtered for removal of low quality reads and adapter trimming. Reads less than 10 bp were discarded from analysis. Target of microRNA's were predicted from chickpea cDNA sequences using CleaveLand 4.0 pipeline and the miRNAs predicted in our analysis (Parween et al., 2015; Addo-Quaye et al., 2009). Reads were mapped to cDNA sequences of chickpea

draft2 genome, obtained from a GFF file information. Targets obtained were grouped into 5 categories, namely, 0, 1, 2, 3 and 4. Category 4: presence of a single raw read. Category 3: Presence of more than 1 read with abundance of the reads being less than both maximum the transcripts and the median. Category 2: Presence of more than 1 raw read with abundance of the reads is less than the maximum number of reads mapped on transcript but more than the median. Category 1: presence of more than 1 raw read with abundance is equal to the maximum which has more than 1 maximum. Category 0: more than 1 raw read with abundance equal to maximum on the transcript having only one maximum (Addo-Quaye et al., 2009). Finally, all of the identified potential targets were annotated and characterized using Blast2GO and transcription factors and kinases were identified using iTAK.

Stemloop and qRT-PCR validation

To validate the miRNAs identified using high-throughput sequencing, stem-loop PCR was performed for both novel and conserved microRNAs. Further expression of their corresponding targets were detected using qRT-PCR. Briefly, 2µg total RNA was used for cDNA synthesis of microRNA and target using Takara SMARTScribe cDNA synthesis kit and BioRad iScript cDNA synthesis kit as per manufacturer's protocol respectively. SYBR green master mix is used for Real Time validation. Each PCR was performed with three technical and biological replicates.

Ectopic expression of cat-miRNAs in chickpea nodules

In order to understand the functional role of miRNAs and its effect on chickpea nodulation, four miRNAs were selected and were overexpressed in chickpea roots by hairy root transformation (Mandal and Sinharoy, 2018). The 200 nt sequences upstream and downstream of the mature miRNAs were extracted from the chickpea kabuli genome (https://legumeinfo.org/) were amplified from the genomic DNA and cloned into the pUb-cGFP-dsRED vector downstream of ubiquitin promoter. The construct containing the transgene was used for chickpea root transformation using the *A. rhizogenes* based hairy root transformation. The transformed roots were screened using stereo fluorescence microscope Nikon AZ100 equipped with Nikon digital camera (Nikon digital sight DS-Ri1) (Nikon) for red fluorescence. The chickpea plants with transformed hairy roots overexpressing the desired miRNA precursors were identified by their constitutive DsRed expression and the nodule number was recorded for each of the overexpressing lines in comparison with vector control.

Results

Sequencing of small RNAs in chickpea root nodule tissue

To gain a comprehensive insight into the landscape of miRNAs involved in root nodulation in chickpea, a miRNA library was generated from pooled nodule tissue at different stages. Sequencing resulted in 21,760,971 small RNA reads. After q20 filtering, adapter removal, sorting reads with length between 20-24 nt and mapping to rRNA, tRNA, snRNA and snoRNA, a total of 4,445,569 reads resulting in 1,315,289 unique tags were obtained (**Table 1).** The 24 nt reads were dominant amongst the sequencing reads followed by 21 nt reads (~30% and 23% unique tags respectively) (**Fig. 1A-B**). The reads were mapped to the chickpea kabuli genome to extract the 250 nt flanking regions which were folded to predict the secondary structures. Structures fulfilling all the latest criteria for miRNA annotation were chosen for miR prediction using ShortStack and mirDeep-P.

miRNA identification in chickpea root nodule tissue

A total of 91 miRNAs were identified through a cumulative curation from both mirDeep-P and ShortStack. For prediction of Novel miRNAs, only those passing the stringent criteria for miRNA annotation through ShortStack were retained (**Table 2**). The secondary stem-loop structures of all the miRNAs were analysed carefully following the latest criteria of miRNA annotation (Axtell and Meyers, 2018) (**Supplementary Fig. 1**). The miRNAs identified were most prevalent in 21 nt category (~69%), followed by 20 nt (16%) (**Fig. 1C**). The first nucleotide at 5' end of miRNA is crucial for biogenesis and processing of miRNAs. Characteristic association specificity of miRNA's with AGO1 and AGO4 is provided by the presence of either U or A as the first nucleotide (O'Brien et al., 2018). It was found that Uridine was the most abundant nucleotide at the first position of 5'end (~78%) in comparison to the rest three nucleotides (**Table 3**). Further, miRNA's with nucleotide length 20, 21, and 22 all had a greater fraction of U as the first nucleotide (93.3%, 73% and 91.6% respectively) and these particularly associate with AGO1 protein (**Table 3**). In contrast, miRNA with nucleotide length 24 has A as first nucleotide and was reported to associate with AGO4 (**Table 3**). The 24 nt miRNAs are reported to mediate DNA methylation through RNA-directed DNA methylation (RdDM) pathway (Yu, 2015).

Out of 91 miRNAs identified, 84 were conserved and remaining 7 were novel. The 84 conserved miRNAs belonged to 24 miRNA families with maximum representation of

miR156 family (~14%) followed by miR166 (8.8%) (**Fig. 1D**). Eighteen of the miRNAs were found to be legume specific of which 13 were found to be chickpea specific (**Table 2**). Though miR159 and the miR166 were the most abundant miRNAs among the conserved miRNAs, miR319 and miR1511 were most prevalent among the legume specific conserved entities (**Table 2**). Amongst the lowest present miRNAs, miR172 and miR164 were the least in the conserved miRNA pool and miR171 among legume specific conserved ones. Chickpea specific miRNAs showed relatively low expression.

The miRNAs identified were located on all 8 chromosomes with maximum miRs being positioned on chromosomes 4 and 6 whereas minimum were located on chromosome 1 (**Fig. 1E**).

Nine pairs of polycistronic miRNAs were identified within 10 kb genomic loci region (**Table 2**). These included miR156, 166, 171, 172, 2111, 396 and miR398 which represented ~21% of the total miRNAs which is in accordance with earlier reported loci in *Populus trichocarpa*, *O. sativa* and Arabidopsis (20%, 19% and 22% respectively) (Merchan et al., 2009).

Phylogenetic analysis of precursor miRNAs in chickpea

To gain an insight into the phylogenetic relationship among the 91 miRNAs identified, a neighbour joining phylogenetic tree was constructed using the precursor sequences of all the miRNAs using MEGA (Kumar et al., 2018) (**Fig. 2**). The tree revealed distribution of precursors sequences into 22 different clades and interestingly, cat-miR482 did not cluster with any of the miRNA precursor sequences. In most cases, miRNAs of the same family clustered in the same clade. However, there were exceptions and several interesting pairs were identified such as cat-MIR408/156c-2, cat-MIR399/1509, cat-NovMIR4/164b, cat-Nov-MIR6/172c/319/394, cat-MIR162a/166b, cat-MIR162b/171f and cat-MIR1507/2118 etc., which were phylogenetically closer to other miRNA instead of their own family member (**Fig. 2**).

Mapping reads on different plant and bacterial genomes

The reads obtained from nodule libraries were mapped on to the genome of chickpea, other legumes, Arabidopsis and rice as mentioned above in materials and method. Reads mapped with different percentage on to different plant genomes with maximum reads (84%) mapping on to chickpea followed by the closest relative *Medicago* (53%) and minimum on the

genomes of Arabidopsis (32%) and rice (29%), the distant species (**Supplementary Fig. 2A**). Similarly the reads obtained from chickpea nodule library were also mapped on bacterial genomes which resulted in maximum reads being mapped on *Mesorhizobium ciceri (CC1192)* genome (13%) and least on *E. coli (strain K12 substr. MG1655)* genome (3%). Similarly, reads from a related study comprising of nodule miRNA from common bean mapped maximally on its host bacterium *Rhizobium tropici (CIAT 899)* (5%) and minimum again on *E. coli* genome (0.4%) (**Supplementary Fig. 2B**) thereby validating the chickpea result.

Multiple sequence alignments and miRNA processing

The chickpea miRNA reads were mapped to these different plant genomes (see materials and method) miRNAs were predicted, from these plant species (**Table 4**). The precursor sequences of the legume for similar mature miRNAs were aligned using MUSCLE and was used to predict the type of processing for miRNA biogenesis. Briefly, three types of processing was predicted to occur during miRNA biogenesis. The base to loop miRNA processing involves recognition of ~15-17 nt conserved stem below the miRNA/miRNA* and first cut by DCL1 below the miRNA/miRNA* region and second cut ~21 nt away from the first cleavage site (Bologna et al., 2013; Chorostecki et al., 2017). The second type was loop to base processing that involves ~15-17 nt conserved region above the miRNA/miRNA* inflicting first cut below the terminal loop region and next cleavage at the base (Bologna et al., 2013; Chorostecki et al., 2017). The third type processing involves, a sub-type of loop to base or base to loop processing with multiple cuts instead of two cuts mentioned above. This results in a sequential loop to base or sequential base to loop processing (Bologna et al., 2013; Chorostecki et al., 2017). The fourth type of miRNA processing during biogenesis includes mixed processing pattern in which there is a short ~6 nt conserved region above and ~4 nt conserved region below the miRNA/miRNA* duplex (Chorostecki et al., 2017). In order to understand the biogenesis mode of miRNAs, precursor sequences of chickpea and other legumes obtained in our analysis were aligned. The alignment of precursors from legumes revealed that cat-miR160, cat-miR162b and cat-miR171f undergoes short loop to base processing (**Fig. 3**) and cat-miR167a, b, c, d, cat-miR168, cat-miR390a and cat-miR396a, b undergoes short base to loop processing (**Fig. 4**). Interestingly, cat-miR166a displays the conserved ~4 nt sequences below the miRNA/miRNA* duplex and other conservations were missing. Therefore, it undergoes mixed processing pattern. Additionally cat-miR319c undergoes sequential loop to base processing involving multiple sequential cuts (**Fig. 5**).

Bacterial small RNA prediction

There is growing evidence, such as the recent study in soybean, to show that the rhizobial tRNA derived small RNAs regulate plant genes during nodulation process using the host Argonaute1 machinery in soybean (Ren et al., 2019). Following the same line of evidence, the chickpea miRNA reads not mapping to the chickpea genome but mapping onto the *M. ciceri* bacterial genome were used to predict small RNAs from *M. ciceri*. Fifty three bacterial small RNAs were predicted using mirDeep-P (**Table 5**). Of these small RNAs, 11 precursor sequences of the bacterial small RNAs were PCR amplified from bacterial RNA (**Supplementary Fig. 3A**). Further, qRT-PCR expression analysis of the smRNAs predicted from bacteria was done using RNA obtained from bacterial cells and nodule tissue. The expression analysis revealed very high expression of the smRNA in nodule tissue compared to the bacterial tissue indicating their presence in root nodule (**Supplementary Fig. 3B**).

In silico expression analysis and identification of chickpea nodule specific miRNAs

To identify root nodule specific miRNAs, a previous report involving identification of miRNAs from different chickpea tissues was used for comparison (Jain et al., 2014). Of the total 91 miRNAs 74 miRNAs were present in all the tissues and 7 were found to be nodule specific (6 conserved miRNAs and 1 novel) (**Fig. 6, Table 6**). 79, 79, 83, 79, 81, 82 and 84 miRNAs were found to be expressed in the tissues such as flower-bud, flower, leaf, young-pod, root, stem and shoot respectively (**Fig. 6, Table 6**). In nodule tissue a total of 16 (14 conserved and 2 novel) miRNAs were found to be highly expressing in comparison to other tissues (**Fig. 6, Table 7**). The miRNAs, especially from miR166 family were found to be the highest expressing in nodule tissue in comparison to others (**Fig. 6, Table 7**). Interestingly, the comparison of root and nodule miRNAs revealed that of the 91 miRNAs, 67 miRNAs were downregulated in nodule tissue in comparison to root tissue whereas 24 were upregulated (**Fig. 6, Table 8**). Twenty eight miRNAs were downregulated more than 10 fold in nodule tissue whereas only 11 were upregulated (10 specific to nodule) more than 10 fold in comparison to root tissue (**Fig. 6, Table 8**).

Degradome sequencing analysis

The miRNA prediction data becomes biologically more relevant and significant only after identification of the corresponding mRNA target transcripts. However, targets predicted from databases such as the psRNA do not provide an experimental validation. Therefore, in this

study, to simultaneously identify and validate targets at global level, high throughput degradome sequencing of four different libraries from uninfected root and *M. ciceri* infected root at various stages (see materials and method) was performed. After quality filtering and adapters trimming, a total of 5,0,517,249, 4,3,578,672, 4,3,870,856 and 5,2,803,247 reads were obtained from the 4 libraries respectively. The CleaveLand pipeline was used to predict targets of the 91 miRs identified from nodules in this study. The abundance of cleavage tags at the target site determined the category of the cleavage targets. They were divided into five categories from 0 to 4 and the maximum number of targets were found in category 2 in all libraries (**Supplementary Fig. 4A**). More than one raw read was present in categories 0, 1, 2 and 3 whereas category 4 had presence of only single read at the cleavage position. The targets obtained from each library were sorted based on p-value < 0.05, Category < 3 and Allen score < 5. Target plots representative of cleavage events were also obtained (**Supplementary Fig. 4C-F)**. In total 351, 387, 421 and 339 cleavage sites were identified from early nodule, mid-late nodule, nodule and root samples that were being targeted by 68, 72, 70 and 67 miRNAs respectively (**Supplementary Fig. 4B, Table 9A-D**). These targeted transcripts were part of the non-redundant set of 564 genes in total, which were cleaved by 85 miRNAs out of the 91 identified. The total sites were predicted to be cleaved by miRNAs were found in 331, 362, 395 and 319 genes from early, mid-late, nodule and root respectively (**Supplementary Fig. 4B, Table 9A-D**). 165 transcripts and 54 miRNAs were present in all the libraries thereby providing strong evidence for their being significant nodulation related miRNAs-target pairs (**Fig. 7A-B**). 245 transcripts were found to be specifically cleaved in nodule libraries in comparison to root and represented nodule specific targets. Similarly, 16 miRNAs were found to be specifically targeting transcripts in nodule tissue only (**Fig. 7A-B**). The conserved miR156 family targeted maximum number of transcripts in each library (~60) and miR1511 had only one target in each library (**Table 9A-D**). Some of the miRNA-target pairs predicted were miR319-*TCP*, miR156-*SPL*, miR164-*NAC*, miR171-*GRAS* and miR172-*AP2* etc., (**Table 9A-D**) which have been previously very well characterized and hence validated the results obtained from the PARE library sequencing.

The target genes were subjected to Blast2GO and iTAK annotation which classified them into the following important categories: transferases (Aspartate aminotransferase, Histone-lysine N-methyltransferase etc.), kinases (adenylate kinase, Serine/threonine-protein kinase 25, Somatic embryogenesis receptor kinase 3, PTI1-like tyrosine-protein kinase 6, Nodulation receptor kinase (NORK), CBL-interacting protein kinase 2, Probable receptor-like protein

kinase, Probable leucine-rich repeat receptor-like serine/threonine-protein kinase, etc.), transcription factors and their associated subunits (TCP, ARF, bHLH, ZF-HD, ZF-CCCH, PPR repeat, TFIIIB, WRKY, ERF, NSP, WD-repeat, Homeoboxleucine zipper protein, RING/FYVE/PHD Zinc finger, NAC, BTB/POZ and TAZ, F-Box, Myb family, heat stress related TF etc.) resistance protein (Disease resistance protein, Disease resistance protein (TIR-NBS class), LRR and NB-ARC domains-containing disease resistance protein etc.) transporters (Proton-coupled amino acid transporter 4, ABC transporter, High-affinity nitrate transporter 3.3, Ammonium transporter 1 member 1, Auxin transporter-like protein 6 etc.), cell division cycle (Cell division cycle protein 48 homolog), ribosomal protein (40s and 60s) and other structural and functional proteins (**Table 9A-D**).

Gene Ontology (GO) and KEGG Analysis of Target Genes

The miRNA targets identified using the PARE libraries were subjected to functional annotation by Blast2GO. The target genes were mainly predicted to function in cellular components like on the basis of their localization in nucleus, cytoplasm, mitocondria, and vesicle during nodule and root development (**Supplementary Fig. 5A**). The miRNA target genes were also predicted to be involved in various biological processes such as protein transport and modification, RNA metabolic process and amide transport (**Supplementary Fig. 5B**). Notably, target genes were also involved in molecular processes such as nucleic acid binding and enzyme activity in both nodule and root development. Apart from binding property, target genes present in root and nodule were also involved in hydrolase activity (**Supplementary Fig. 5C**).

The KEGG pathway analysis revealed an involvement of the cleaved targets in 284 pathways of which 98 had common transcripts from all the 4 libraries (**Fig. 7C**). Further, genes present in 6 pathways were associated with root tissue only whereas 110 transcripts were involved in pathways associated only with nodulation process (**Fig. 7C**). Some of the important pathways include calcium signaling pathway (**Supplementary Fig. 6A**), Nod-like receptor signaling pathway (**Supplementary Fig. 6B**), plant hormone signal transduction (**Supplementary Fig. 6C**) and plant pathogen interaction (**Supplementary Fig. 6D**).

Expression profiling of conserved and novel miRs and their corresponding targets

In order to check the abundance of miRNA transcripts across chickpea tissues, a stem-loop qRT-PCR was performed. Expression analysis was carried out for 8 conserved miRNAs

(having indicative role in nodule development) and the 7 novel miRNAs. Expression analysis revealed that pattern of miRNA, stem-loop qRT-PCR was performed. Expression profiles of all the novel miRNAs and 8 conserved miRNAs were performed. The expression profile revealed that cat-miR1509, 398a-5p, cat-NovmiR1, 2, 4 and cat-NovmiR5 showed fairly high expression during nodule process (**Fig. 8A**).

The miRNAs are known to target genes for degradation and hence are expected to exhibit expression pattern, which are antagonistic to their targets. Therefore, to validate the targets predicted from PARE libraries and their negative correlation with the miRNAs, qRT-PCR was employed. The abundance of miR166a, 171c, 172c, 390, 394 and NovmiR1 were antagonistic with its target *IAA, NSP2, AP2, PTI1, HP1* and *Annexin* respectively. Additionally, expression of miR171f and its targets *SCL-6* and *NORK*, were found to be antagonistic. The legume specific miR1509 targeted transcripts of *AK* and *Jason*, and their expression were anti-correlated. (**Fig. 8B-I**)

Ectopic overexpression of cat-miR171, cat-miR172, cat-miR394 and cat-miR1509 by hairy root transformation

The NGS sequencing and analysis of miRNA and PARE libraries yielded several important miRNA-target pairs. Using this analysis, four miRNAs, cat-miRs171f, cat-miR172c, cat-miR394 and cat-miR1509 were selected based on their significance as described above. The ectopically overexpressing lines of these miRNAs were obtained using hairy root transformation (**Fig. 9A**).To establish the effect of miRNAs activity on their respective targets qRT-PCR based expression levels of miRNAs and corresponding targets were analysed in the overexpression lines. The overexpression lines of cat-miRs171f, displayed more than fivefold increase in cat-miRs171f levels with a concomitant decrease in targets *SCL-6* and *NORK* to 0.67 and 0.55 in comparison to vector control (**Fig. 10A**). The effect of the cat-miRs171f overexpression and downregulation of corresponding target resulted in a significant decrease in nodule number to 0.71 times in comparison to vector control (**Fig. 9B**).

Similarly, chickpea roots transformed with overexpression construct of miR172c displayed enhancement of the miRNA upto ~15-fold in comparison to the vector control with subsequent reduction of its target AP2 transcript by 0.37 times (**Fig. 10B**). This ultimately led to an increase in nodule number by 1.67 fold (**Fig. 9B**).

The role of miR394 in nodule symbiosis remains largely unexplored till date. Therefore, to investigate its role in chickpea nodulation ectopic expression of cat-miR394 was carried out. Analysis of the overexpression lines showed a 1.48 times increase in nodule number (**Fig. 9B)**. This could be explained by the observed 4-fold increase in expression of cat-miR394 coupled with a 0.65 fold decrease in expression of the target *HP1* gene (**Fig. 10C)**.

Likewise the legume specific miRNA, cat-miR1509 was also analysed for its *in-vivo* effect on the nodulation process. The overexpression of cat-miR1509 resulted in a 2.07 fold increase in the expression of the miR and a concomitant decrease in the expression of target *AK* transcript by 0.67 times corresponding to the vector control (**Fig. 10D**). The overexpression of cat-miR1509 also resulted in an increase in nodule number by 1.56 times of vector control (**Fig. 9B)**.

Discussion

The crop legume is not only valued for its protein content, but for the biological nitrogen fixation (BNF) that it can bring about with the help of rhizobia in its nodules. Reports of spatial and temporal expression of miRNAs are available in chickpea (Jain et al., 2014; Kohli et al., 2014; Khandal et al., 2017; Tiwari and Bhatia, 2019). However, miRNA identification and functional characterization has not been performed in crop chickpea nodules. Hence, for identification of miRNAs involved in chickpea nodulation, a high-throughput NGS based sequencing was carried out. A small RNA library from nodules was sequenced and analysed resulting in identification of 84 conserved and 7 novel miRNAs. Moreover as expected from a chickpea nodule library, legume, chickpea and nodule specific miRNAs were also identified. The first nucleotide of the miRNAs determines the biogenesis and final processivity. Both strands of the miRNA duplex can be loaded into the Argonaute (AGO). The U at the first nucleotide position at the 5’ end of the miRNA duplex decreases the thermodynamic stability making it more favourable to be loaded into AGO as the guide strand (O’Brien et al., 2018). In our study 71 miRs had U as the first nucleotide at 5’ end and were hence amenable to DCL1 cleavage and AGO1 association. The 21-nt miRNAs were the most predominant miRNAs in this study however, interestingly 12 miRNAs were also identified with 22-nt length. Such miRNAs are reported to trigger the biogenesis of secondary siRNAs from target mRNAs (Chen et al., 2010). Our analysis identified novel polycistronic precursors with non-homologous miRNA (miR156, miR166, miR171, miR172, miR2111, miR396 and miR398).

These polycistron harbouring miRNAs target functionally related genes and hence can be used as a tool to regulate gene expression (Merchan et al., 2009).

To evaluate the genetic relatedness of the miRNAs, a phylogenetic tree based on the precursor miRNA sequences was constructed. Each subclade contained miRNA members of the same family, however in some cases, very distinct unrelated miRNAs clustered closely together. This close relationship indicated a common ancestor gene that later underwent functional diversification to enable cleavage of different targets. The miRNA reads were mapped onto genomes of legumes and non-legumes. The percentage of reads mapped was in accordance with the evolutionary distance of the organisms from each other indicating that the same pattern of evolution was followed by the small RNAs as other genes. Similarly the abundance of reads mapped onto bacterial genomes also followed similar pattern indicating a co-evolution of the rhizobia along with the host legumes. Further, the miRNA reads mapping to the *M. ciceri* genome were used to predict bacterial small RNA. Recently it has been shown that rhizobial tRNA derived small RNAs regulate the plant genes governing nodule numbers (Ren et al., 2019). Our data of bacterial smRNA prediction and their expression validation opens up several new areas for investigation including the cross kingdom biogenesis and translocation of small RNAs.

An evolutionary footprint analysis of the precursor miRNA alignments showed that there were conserved regions other than miRNA/miRNA* present in the plant precursors. Their relative positions and length vary among different miRNAs, which determines the processing mechanism of miRNAs. An alignment analysis of precursor sequences from miRNAs identified in chickpea and other legumes revealed regions with conserved sequences beyond the miRNA/miRNA* and these evolutionary footprints were linked to mechanistic processing during miRNA biogenesis. The findings of this investigation revealed that the conserved processing mechanism of the precursor cannot be uncoupled from the miRNA sequence and the evolutionary footprints that determine the biogenesis (Bologna et al., 2013; Chorostecki et al., 2017).

The *in-silico* expression analysis of the miRNAs indicated a spatio-temporal expression pattern of the miRNAs with several miRNAs recruited to specifically regulate the plant developmental processes. Expression analysis of the cat-miRs revealed 7 to be nodule specific (cat-miR156f-1, 156f-2, 162a, 162b, 168, 171c and cat-NovmiR7). The nodule specific cat-miR171c targets *NSP2* gene, which is an important TF involved in nodulation.

Overexpression of cat-miR172c in this study led to a significant decrease in nodule number. Following this, the other six miRNAs could also regulate expression of key targets regulating nodulation and could be interesting candidates to be characterized for their role during nodulation process. Of the 16 highly expressing miRNAs in nodule tissue, the important member miR166 has been shown to have a polycistonic precursor. In the model legume *Medicago truncatula*, a single transcriptional unit, the MtMIR166a precursor consisted of tandem copies of miR166. The miR166 family regulates a nodule development related TF family, the class-III homeodomain-leucine zipper (HD-ZIP III) genes. *In-planta* overexpression of miR166 causes, miR166 levels to accumulate leading to downregulation of several class-III HD-ZIP genes and a reduction of the number of nodules and lateral roots in the transgenics (Boualem et al., 2008). This study established that even though the plant polycistronic miRNA precursors, are rare, they can be processed to regulate the genes and associated pathways. Out study also identified the polycistronic cat-miR166 and its family members that showed high expression in chickpea nodule tissue. Hence it could be suggested that a similar mechanism for processing of polycistronic miRNAs was in place in chickpea nodules also.

To lend biological significance to the study of miRNA identification, their target prediction becomes necessary. In this study analysis of the 4 PARE libraries identified several significant miRNA-target pairs, some of which were earlier well characterized for their tole in nodulation. This gave us the confidence to pick up other novel miRNA-target pairs identified in this study as candidates for analysis. Our study showed that 165 targets of 54 miRNAs were found to occur in all the 4 PARE libraries and could serve as high confidence pairs for further characterization. From the venn diagram analysis (**Fig. 7A-C**) of the cleaved targets and targeting miRNAs it could be deduced that the 12 cleaved targets were likely to be targeted by the 5 miRNAs regulating 7 functional pathways commonly found in all 4 libraries.

In the past decade, the miRNA mediated regulation has been well established in several plant species including characterization of miRNAs regulating symbiosis at different stages (Boualem et al., 2008; De Luis et al., 2012; Nizampatnam et al., 2015; Nova-Franco et al., 2015). However, functional characterization of a miRNA involved in nodulation in the important crop plant chickpea has never been done. Therefore in order to functionally decode

the regulatory circuitry of miRNA and their targets, 4 miRNA candidates were selected for functional characterization based on the importance of their targets in nodulation.

The miR171 family has been documented to negatively impact the nodulation process and mycorrhizal symbiosis by targeting *NSP2* gene and *SCL-6* (De Luis et al., 2012; Lauressergues et al., 2012; Hofferek et al., 2014; Hossain et al., 2019). The ectopic expression of gma-miR171o and gma-miR171q resulted in a significant inhibition of the symbiosis with very few nodules devoid of nitrogen fixation (Hossain et al., 2019). In *M. truncatula* it has been reported that cytokinin regulates *NSP2* post-transcriptionally during nodule organogenesis (Ariel et al., 2012) through miR171h. The PARE library analysis in our study revealed two targets of cat-miR171f, i.e. *SCL-6* and *NORK*. Through transient *in planta* expression in chickpea, the cleavage of the two targets by cat-miR171f was confirmed. The expression levels of both the targets displayed negative correlation with the cat-miR171f in the overexpression lines. Additionally the overexpression lines resulted in a significant reduction in nodule numbers. *SCL-6* belongs to the GRAS family of transcription factor which are known to regulate the shoot and root radial patterning (Lim et al., 2000). It has been reported in soybean that the miR171 targeting *SCL-6* reduced the nodule numbers but role of *SCL-6* in nodulation is yet to be characterized. Similarly the other target *NORK*, which is a LRR-RLK has been reported to function in Nod-factor perception and downstream transduction system that initiates a signal cascade leading to nodulation (Endre et al., 2002). The ortholog of *NORK*, *DMI2* has been reported to act downstream of Nod-factor receptors and is activated by rhizobial inoculation (Pan et al., 2018). Overexpression of full length protein or the kinase domain can induce nodulation even in absence of bacteria (Antolín-Llovera et al., 2014). Therefore, reduction in nodule number in the overexpression lines of cat-miR171f were amply supported by previous literature.

The miR172 is well known for its role in flowering time and phase transition process, tuberization of potato and also recently its role in nodule symbiosis has been elucidated (Zhu and Helliwell, 2011; Martin et al., 2009; Nova-Franco et al., 2015; Holt et al., 2015; Yan et al., 2013; Wang et al., 2014). The miR172 is known to target AP2 family member named *nodule number control1* (NNC1). The miR172c-*NNC1* pair acts as a critical regulatory module of symbiosis which regulates *ENOD40* downstream which leads to nodule initiation and organogenesis (Wang et al., 2014). Recently it has also been established that miR172c/*NNC1* module directly controls the transcription of *GmRIC1* and *GmRIC2*, (the cle peptides) which mediate AON activation and attenuation. Based on the importance, cat-

miR172-c was selected for overexpression analysis to validate its role in chickpea. Five members of the cat-miR172 miRNA were identified in our study of which miR172c specifically targets the AP2 TF superfamily. The target AP2 transcript was identified through PARE library analysis. Overexpression of cat-miR172c resulted in an increase in nodule numbers which could be explained by the cat-miR172c-NNCI mediated ENOD activation.

The analysis of miRNA-PARE libraries revealed cat-miR394 to target an important cytokinin signaling molecule *histidine phosphotransferase* (HP) with a significant p-value in all the PARE libraries. This target (*HP1*) of cat-miR394 was a novel target reported in this study for the first time. HPs are important regulatory molecules which relay the transfer of the phosphate group from the histidine kinases (HKs) to response regulators (RRs) (Hwang et al., 2002). Although the mechanism by which HPs regulate the symbiosis process has not been elucidated. The *in-planta* expression studies overexpressing cat-miR394 revealed a significant phenotype (increase in nodule number). The mechanism underlying the cat-miR394-*HP1* network mediated nodulation need to be further addressed.

The legume-specific cat-miR1509 targets *AK* and *Jason* as revealed by the degradome analysis. (convert ADP to ATP) and AMP (2ADP ↔ ATP + AMP). The enzyme is known to regulate the level of phosphorus and starch (Regierer et al., 2002; Carrari et al., 2005). There are no reports available currently to decipher the role of *AK* during nodulation. In present study the cat-miR1509 which target *AK* was overexpressed and resulted in an increase in nodule numbers. This result was indicative of the role of cat-miR1509 mediated cleavage of *AK* and hence a phenotype change in form of increased nodulation. The other target *Jason* did not showed an antagonistically correlated expression with the miRNA in the cat-miR1509 overexpression lines. The regulation of nodule phenotype by miR1509 and *AK* needs to be deal with a deeper insight in future.

Overall the following study presents a holistic view of the miRNAs governing the post-transcriptional gene regulation in chickpea nodulation. A miRNA and 4 PARE libraries were sequenced from the nodule tissues leading to the identification of a repertoire of conserved and the novel miRNAs and their targets implicated to have a role in nodulation. The miRNAs identified were extensively studied for their phylogenetic relationships, prediction of biogenesis from precursors, and their *in-silico* expression patterns. Further, the reads were mapped onto rhizobia genome to predict the bacterial small RNAs with putative cross-kingdom roles that need to be investigated. Based on the significance of the identified

miRNAs and targets involved in nodulation, 4 miRNAs were chosen for functional *in-planta* studies. This revealed a novel target, *NORK,* an ortholog of *DMI2* belonging to the LRR-RLK family. The *NORK* was targeted by miR171f and the overexpression of the miR suppressed the nodule number. The miR171f is a cytokinin regulated miRNA and hence this miR and its target pair (a LRR-RLK) explains the effect played by cytokinin on nodulation as proposed earlier. Additionally, a well characterized target of miR172c has also been functionally validated that showed an increase in nodule number as has been previously reported. Two conserved miRNAs (cat-miR394, cat-miR1509) which have not been previously assigned any role during nodulation, nor have their target genes been characterized, were also chosen for investigation of their *in-vivo* role in chickpea nodulation. Their overexpression resulted in an increase in nodule number. The present study provides several novel clues about the conserved and the novel miRNAs, bacterial small RNAs and important nodulation related unexplored target genes. Further, the miRNAs characterized in the present study could be investigated in detail to understand their involvement in regulatory pathways and the mechanism of their action. The genome-wide resources generated in the present study would serve as a foundation to be can be utilized by the research community to enhance the understanding related to modulation of PTGS by miRNAs during symbiosis.

References

Addo-Quaye, C., Miller, W., and Axtell, M. (2009). CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics **25**.

Antolín-Llovera, M., Ried, M.K., and Parniske, M. (2014). Cleavage of the SYMBIOSIS RECEPTOR-LIKE KINASE Ectodomain Promotes Complex Formation with Nod Factor Receptor 5. Curr. Biol. **24**: 422–427.

Ariel, F. et al. (2012). Two Direct Targets of Cytokinin Signaling Regulate Symbiotic Nodulation in Medicago truncatula. Plant Cell **24**: 3838–3852.

Axtell, M.J. (2013). ShortStack: Comprehensive annotation and quantification of small RNA genes. Rna **19**: 740–751.

Axtell, M.J. and Meyers, B.C. (2018). Revisiting Criteria for Plant MicroRNA Annotation in the Era of Big Data. Plant Cell **30**: 272–284.

Bartel, D.P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell

116: 281–297.

van den Berg, A., Mols, J., and Han, J. (2008). RISC-target interaction: cleavage and translational suppression. Biochim. Biophys. Acta **1779**: 668–677.

Bologna, N.G., Schapire, A.L., Zhai, J., Chorostecki, U., Boisbouvier, J., Meyers, B.C., and Palatnik, J.F. (2013). Multiple RNA recognition patterns during microRNA biogenesis in plants. Genome Res. **23**: 1675–1689.

Boualem, A., Laporte, P., Jovanovic, M., Laffont, C., Plet, J., Combier, J.-P., Niebel, A., Crespi, M., and Frugier, F. (2008). MicroRNA166 controls root and nodule development in Medicago truncatula. Plant J. **54**: 876–87.

Carrari, F., Coll-Garcia, D., Schauer, N., Lytovchenko, A., Palacios-Rojas, N., Balbo, I., Rosso, M., and Fernie, A.R. (2005). Deficiency of a plastidial adenylate kinase in Arabidopsis results in elevated photosynthetic amino acid biosynthesis and enhanced growth. Plant Physiol. **137**: 70–82.

Chen, H.-M., Chen, L.-T., Patel, K., Li, Y.-H., Baulcombe, D.C., and Wu, S.-H. (2010). 22-nucleotide RNAs trigger secondary siRNA biogenesis in plants. Proc. Natl. Acad. Sci. **107**: 15269 LP – 15274.

Chorostecki, U., Moro, B., Rojas, A.M.L., Debernardi, J.M., Schapire, A.L., Notredame, C., and Palatnik, J.F. (2017). Evolutionary Footprints Reveal Insights into Plant MicroRNA Biogenesis. Plant Cell **29**: 1248 LP – 1261.

Chuck, G., Candela, H., and Hake, S. (2009). Big impacts by small RNAs in plant development. Curr. Opin. Plant Biol. **12**: 81–6.

Combier, J.-P.P. et al. (2006). MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in Medicago truncatula. Genes Dev. **20**: 3084–8.

Devers, E.A., Branscheid, A., May, P., and Krajinski, F. (2011). Stars and symbiosis: microRNA- and microRNA*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis. Plant Physiol **156**.

Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P., and Kiss, G.B. (2002). A receptor kinase gene regulating symbiotic nodule development. Nature **417**: 962–966.

Formey, D. et al. (2014). The small RNA diversity from Medicago truncatularoots under biotic interactions evidences the environmental plasticity of the miRNAome. Genome Biol. **15**: 457.

Formey, D., Iñiguez, L.P., Peláez, P., Li, Y.-F., Sunkar, R., Sánchez, F., Reyes, J.L., and Hernández, G. (2015). Genome-wide identification of the Phaseolus vulgaris sRNAome using small RNA and degradome sequencing. BMC Genomics **16**: 423.

German, M.A., Luo, S., Schroth, G., Meyers, B.C., and Green, P.J. (2009). Construction of Parallel Analysis of RNA Ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. Nat. Protoc. **4**: 356–362.

Hamilton, A.J. and Baulcombe, D.C. (1999). A Species of Small Antisense RNA in Posttranscriptional Gene Silencing in Plants. Science (80-.). **286**: 950 LP – 952.

Hofferek, V., Mendrinna, A., Gaude, N., Krajinski, F., and Devers, E. a (2014). MiR171h restricts root symbioses and shows like its target NSP2 a complex transcriptional regulation in Medicago truncatula. BMC Plant Biol. **14**: 199.

Holt, D.B., Gupta, V., Meyer, D., Abel, N.B., Andersen, S.U., Stougaard, J., and Markmann, K. (2015). micro RNA 172 (miR172) signals epidermal infection and is expressed in cells primed for bacterial invasion in Lotus japonicus roots and nodules. New Phytol. **208**: 241–256.

Hossain, M.S., Hoang, N.T., Yan, Z., Tóth, K., Meyers, B.C., and Stacey, G. (2019). Characterization of the Spatial and Temporal Expression of Two Soybean miRNAs Identifies SCL6 as a Novel Regulator of Soybean Nodulation . Front. Plant Sci. **10**: 475.

Hwang, I., Chen, H.-C., and Sheen, J. (2002). Two-Component Signal Transduction Pathways in Arabidopsis. Plant Physiol. **129**: 500 LP – 515.

Jain, M., Chevala, V.V.S.N., and Garg, R. (2014). Genome-wide discovery and differential regulation of conserved and novel microRNAs in chickpea via deep sequencing. J. Exp. Bot. **65**: 5945–5958.

Khandal, H., Parween, S., Roy, R., Meena, M.K., and Chattopadhyay, D. (2017). MicroRNA profiling provides insights into post-transcriptional regulation of gene

expression in chickpea root apex under salinity and water deficiency. Sci. Rep. **7**: 4632.

Kohli, D., Joshi, G., Deokar, A.A., Bhardwaj, A.R., Agarwal, M., Katiyar-Agarwal, S., Srinivasan, R., and Jain, P.K. (2014). Identification and Characterization of Wilt and Salt Stress-Responsive MicroRNAs in Chickpea through High-Throughput Sequencing. PLoS One **9**: 1–19.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol. Biol. Evol. **35**: 1547–1549.

Lauressergues, D., Delaux, P.-M., Formey, D., Lelandais-Brière, C., Fort, S., Cottaz, S., Bécard, G., Niebel, A., Roux, C., and Combier, J.-P. (2012). The microRNA miR171h modulates arbuscular mycorrhizal colonization of Medicago truncatula by targeting NSP2. Plant J. **72**: 512–522.

Lelandais-Brière, C., Naya, L., Sallet, E., Calenge, F., Frugier, F., Hartmann, C., Gouzy, J., and Crespi, M. (2009). Genome-wide Medicago truncatula small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. Plant Cell **21**.

Lim, J., Helariutta, Y., Specht, C.D., Jung, J., Sims, L., Bruce, W.B., Diehn, S., and Benfey, P.N. (2000). Molecular analysis of the SCARECROW gene in maize reveals a common basis for radial patterning in diverse meristems. Plant Cell **12**: 1307–1318.

De Luis, A., Markmann, K., Cognat, V., Holt, D.B., Charpentier, M., Parniske, M., Stougaard, J., and Voinnet, O. (2012). Two microRNAs linked to nodule infection and nitrogen-fixing ability in the legume Lotus japonicus. Plant Physiol. **160**: 2137–54.

Mandal, D. and Sinharoy, S. (2018). A Toolbox for Nodule Development Studies in Chickpea: A Hairy-Root Transformation Protocol and an Efficient Laboratory Strain of Mesorhizobium sp. Mol. Plant-Microbe Interact. **32**: 367–378.

Martin, A., Adam, H., Diaz-Mendoza, M., Zurczak, M., Gonzalez-Schain, N.D., and Suarez-Lopez, P. (2009). Graft-transmissible induction of potato tuberization by the microRNA miR172. Development **136**: 2873–2881.

Merchan, F., Boualem, A., Crespi, M., and Frugier, F. (2009). Plant polycistronic

precursors containing non-homologous microRNAs target transcripts encoding functionally related proteins. Genome Biol. **10**: R136–R136.

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. EMBO J. **19**: 5194–5201.

Mishra, N.S. and Mukherjee, S.K. (2007). A Peep into the Plant miRNA World.: 1–9.

Nizampatnam, N.R., Schreier, S.J., Damodaran, S., Adhikari, S., and Subramanian, S. (2015). microRNA160 dictates stage-specific auxin and cytokinin sensitivities and directs soybean nodule development. Plant J. **84**: 140–153.

Nova-Franco, B., Íñiguez, L.P., Valdés-López, O., Alvarado-Affantranger, X., Leija, A., Fuentes, S.I., Ramírez, M., Paul, S., Reyes, J.L., Girard, L., and Hernández, G. (2015). The Micro-RNA172c-APETALA2-1 Node as a Key Regulator of the Common Bean-Rhizobium etli Nitrogen Fixation Symbiosis. Plant Physiol. **168**: 273 LP – 291.

O'Brien, J., Hayder, H., Zayed, Y., and Peng, C. (2018). Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation . Front. Endocrinol. **9**: 402.

Pan, H., Stonoha-Arther, C., and Wang, D. (2018). Medicago Plants Control Nodulation by Regulating Proteolysis of the Receptor-Like Kinase DMI2. Plant Physiol. **177**: 792–802.

Parween, S., Nawaz, K., Roy, R., Pole, A.K., Venkata Suresh, B., Misra, G., Jain, M., Yadav, G., Parida, S.K., Tyagi, A.K., Bhatia, S., and Chattopadhyay, D. (2015). An advanced draft genome assembly of a desi type chickpea (Cicer arietinum L.). Sci. Rep. **5**: 12806.

Regierer, B., Fernie, A.R., Springer, F., Perez-Melis, A., Leisse, A., Koehl, K., Willmitzer, L., Geigenberger, P., and Kossmann, J. (2002). Starch content and yield increase as a result of altering adenylate pools in transgenic plants. Nat. Biotechnol. **20**: 1256–1260.

Ren, B., Wang, X., Duan, J., and Ma, J. (2019). Rhizobial tRNA-derived small RNAs are signal molecules regulating plant nodulation. Science (80-.). **365**: 919–922.

Rogers, K. and Chen, X. (2013). Biogenesis, Turnover, and Mode of Action of Plant MicroRNAs. Plant Cell **25**: 2383–99.

Song, J., Smith, S.K., and Hannon, G.J. (2004). Crystal Structure of Argonaute Slicer Activity. **305**: 1434–1441.

Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W.B., Zhu, J.-K., and Yu, O. (2008). Novel and nodulation-regulated microRNAs in soybean roots. BMC Genomics **9**: 160.

Tiwari, M. and Bhatia, S. (2019). Expression profiling of miRNAs indicates crosstalk between phytohormonal response and rhizobial infection in chickpea. J. Plant Biochem. Biotechnol.

Tiwari, M., Yadav, M., Singh, B., Pandey, V., Nawaz, K., and Bhatia, S. (2019). Global analysis of Two Component System (TCS) member of chickpea and other legume genomes implicates its role in enhanced nodulation. bioRxiv: 667741.

Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. Cell **136**: 669–87.

Wang, Y., Li, P., Cao, X., Wang, X., Zhang, A., and Li, X. (2009). Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. Biochem. Biophys. Res. Commun. **378**: 799–803.

Wang, Y., Wang, L., Zou, Y., Chen, L., Cai, Z., Zhang, S., Zhao, F., Tian, Y., Jiang, Q., Ferguson, B.J., Gresshoff, P.M., and Li, X. (2014). Soybean miR172c Targets the Repressive AP2 Transcription Factor NNC1 to Activate ENOD40 Expression and Regulate Nodule Initiation. Plant Cell **26**: 4782 LP – 4801.

Yan, Z., Hossain, M.S., Wang, J., Valdés-López, O., Liang, Y., Libault, M., Qiu, L., and Stacey, G. (2013). miR172 regulates soybean nodulation. Mol. Plant-Microbe Interact. **26**: 1371–7.

Yang, X. and Li, L. (2018). miRDeep-P□: a computational tool for analyzing the microRNA transcriptome in plants. **27**: 2614–2615.

Yu, M.X. and Bin (2015). siRNA-directed DNA Methylation in Plants. Curr. Genomics **16**: 23–31.

Zhang, B. and Wang, Q. (2014). MicroRNA-based Biotechnology for Plant Improvement. J. Cell. Physiol.

Zhu, Q.-H. and Helliwell, C.A. (2011). Regulation of flowering time and floral patterning by miR172. J. Exp. Bot. **62**: 487–495.

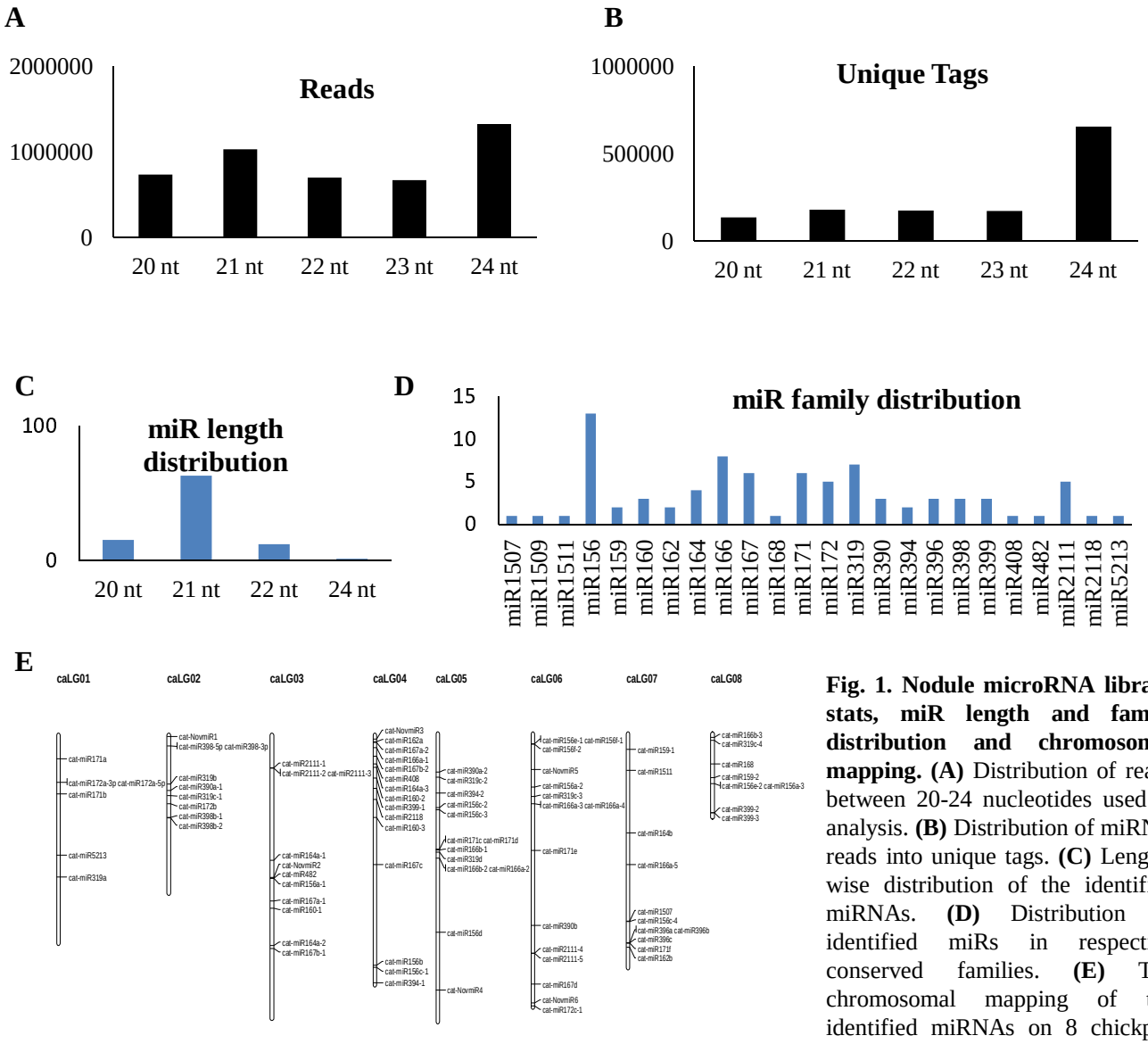


Fig. 1. Nodule microRNA library stats, miR length and family distribution and chromosomal mapping. (A) Distribution of reads between 20-24 nucleotides used in analysis. **(B)** Distribution of miRNA reads into unique tags. **(C)** Length-wise distribution of the identified miRNAs. **(D)** Distribution of identified miRs in respective conserved families. **(E)** The chromosomal mapping of the identified miRNAs on 8 chickpea chromosomes.

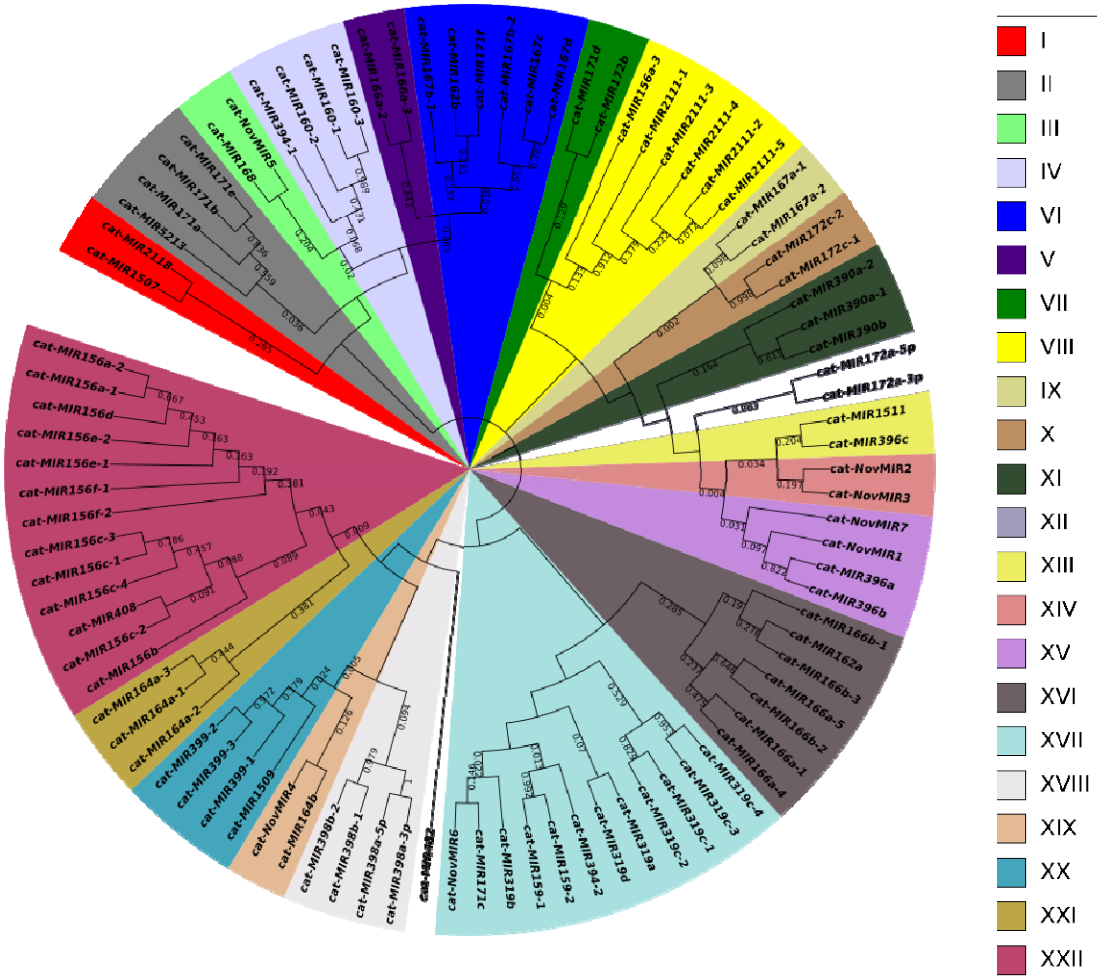


Fig. 2. Phylogenetic analysis of the predicted pre-miRNA sequences from chickpea. The phylogenetic tree revealing clustering of of the 91 miRNA precursor sequences from chickpea in different subclades using MEGA. The bootstrap values are marked as numbers.

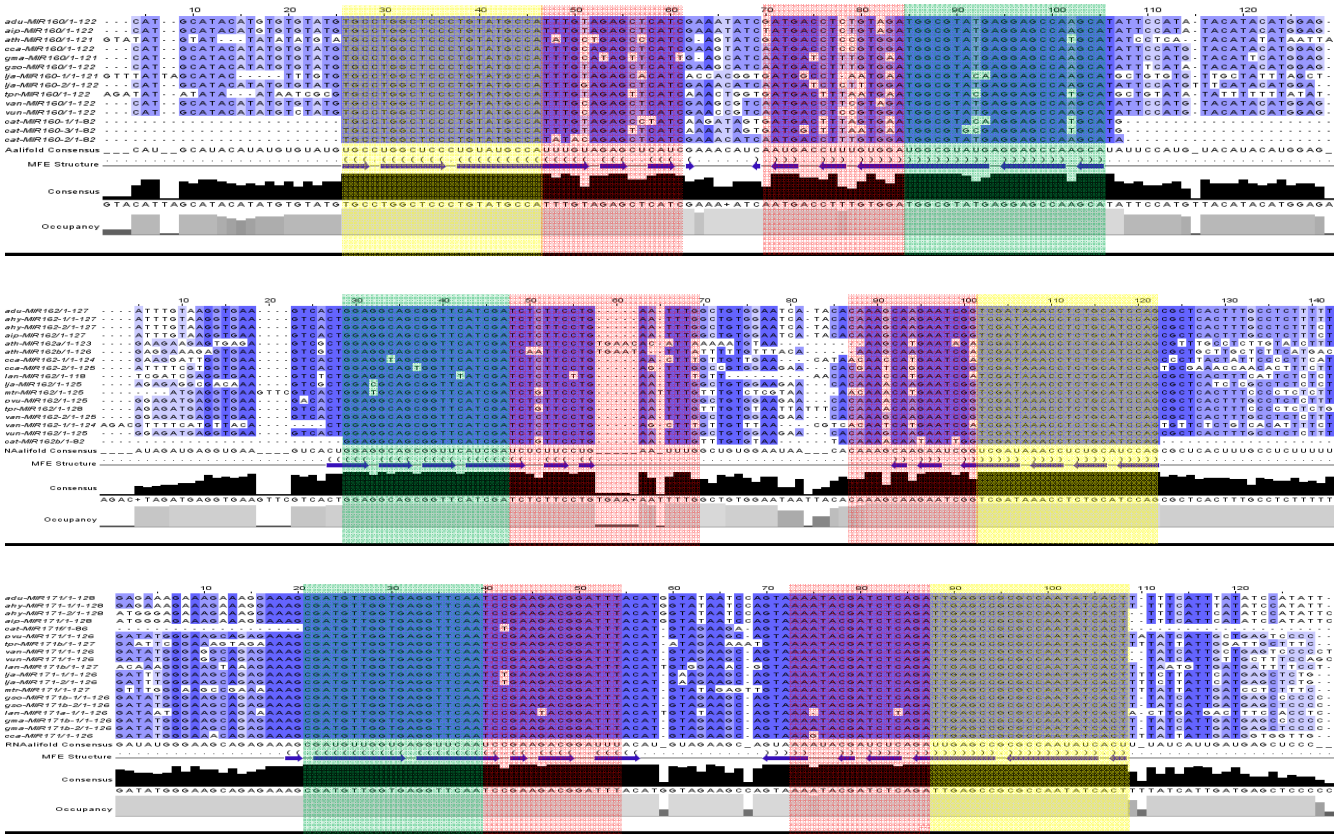


Fig. 3. Representation of miRNA Precursors Processed in a Loop-to-Base Direction. Alignment of precursor sequences of miR160, 162 and171 representing a short loop -to- base processing. Yellow marking indicate mature miRNA sequences, green marking indicate miRNA* sequence. Red marking indicate conserved sequences above miRNA/miRNA*

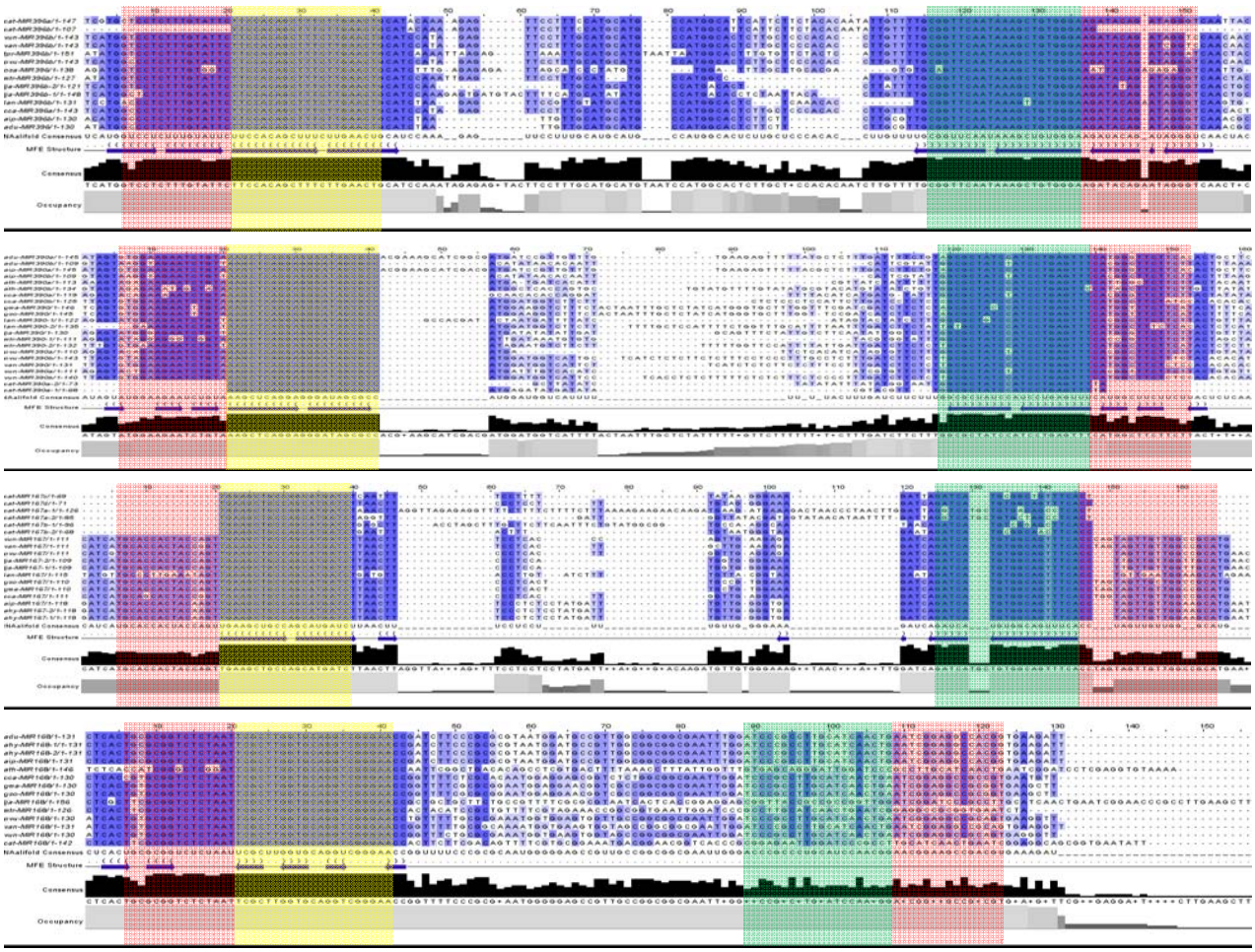


Fig. 4. Representation of miRNA Precursors Processed in a Base-to-Loop Direction. Alignment of precursor sequences of miR167, 168, 390 and 396 representing a short base-to-loop processing. Yellow marking indicate mature miRNA sequences, green marking indicate miRNA* sequence. Red marking indicate conserved sequences above miRNA/miRNA*

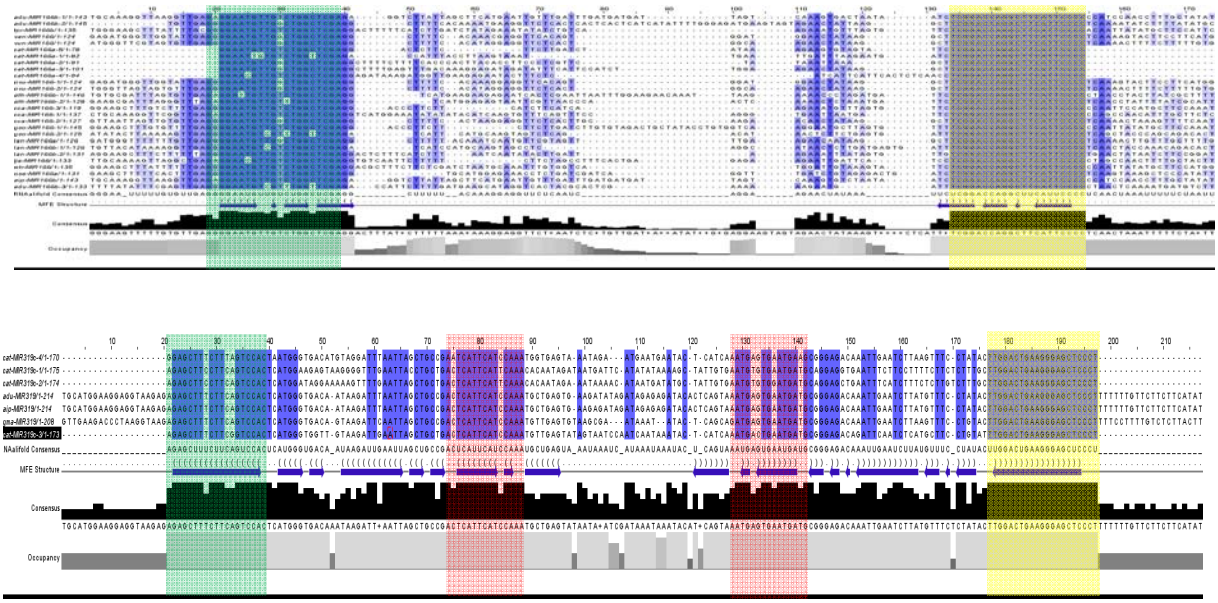


Fig. 5. Representation of miRNA Precursors Processed in a Sequential Loop-to-Base Direction. Alignment of precursor sequences of miR166 (upper panel) representing a mixed processing and miR319 (lowerpanel) representing a sequential loop -to- base processing. Yellow marking indicate mature miRNA sequences, green marking indicate miRNA* sequence. Red marking indicate conserved sequences above miRNA/miRNA*

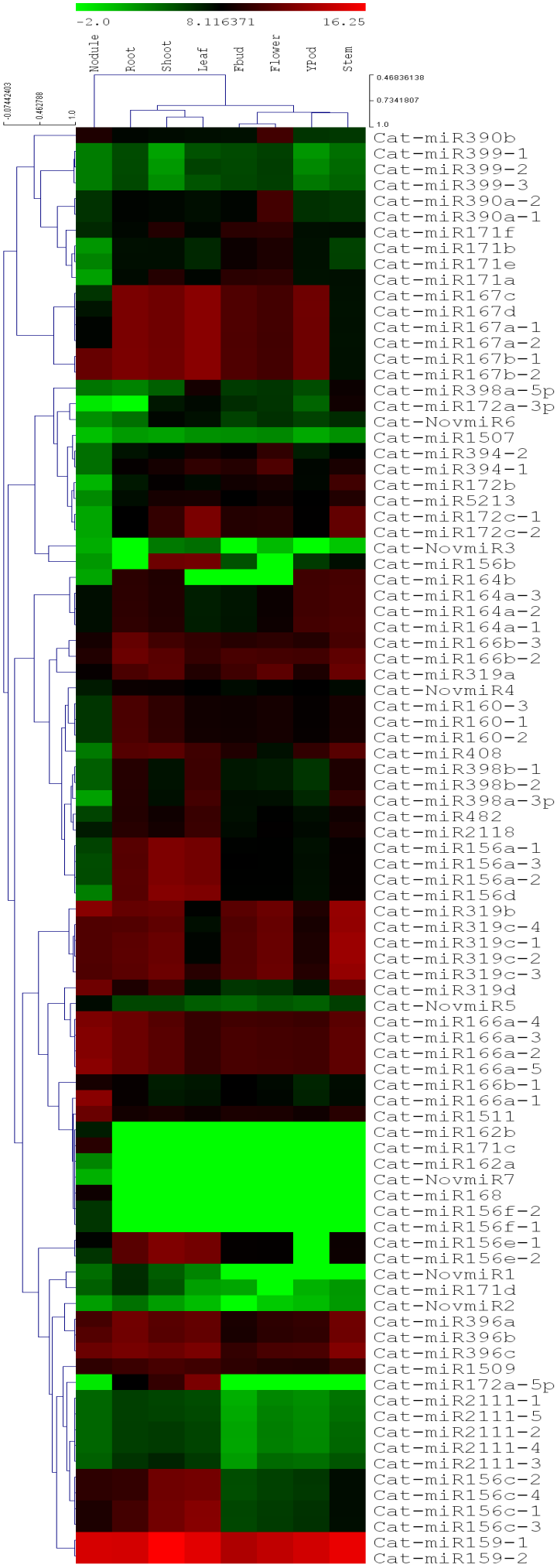


Fig. 6. In-silico expression profile of miRNAs in different chickpea tissues. A, heat-map and 2-way hierarchical clustering based on 91 miRNAs that were differentially expressed between nodule and other chickpea tissues.

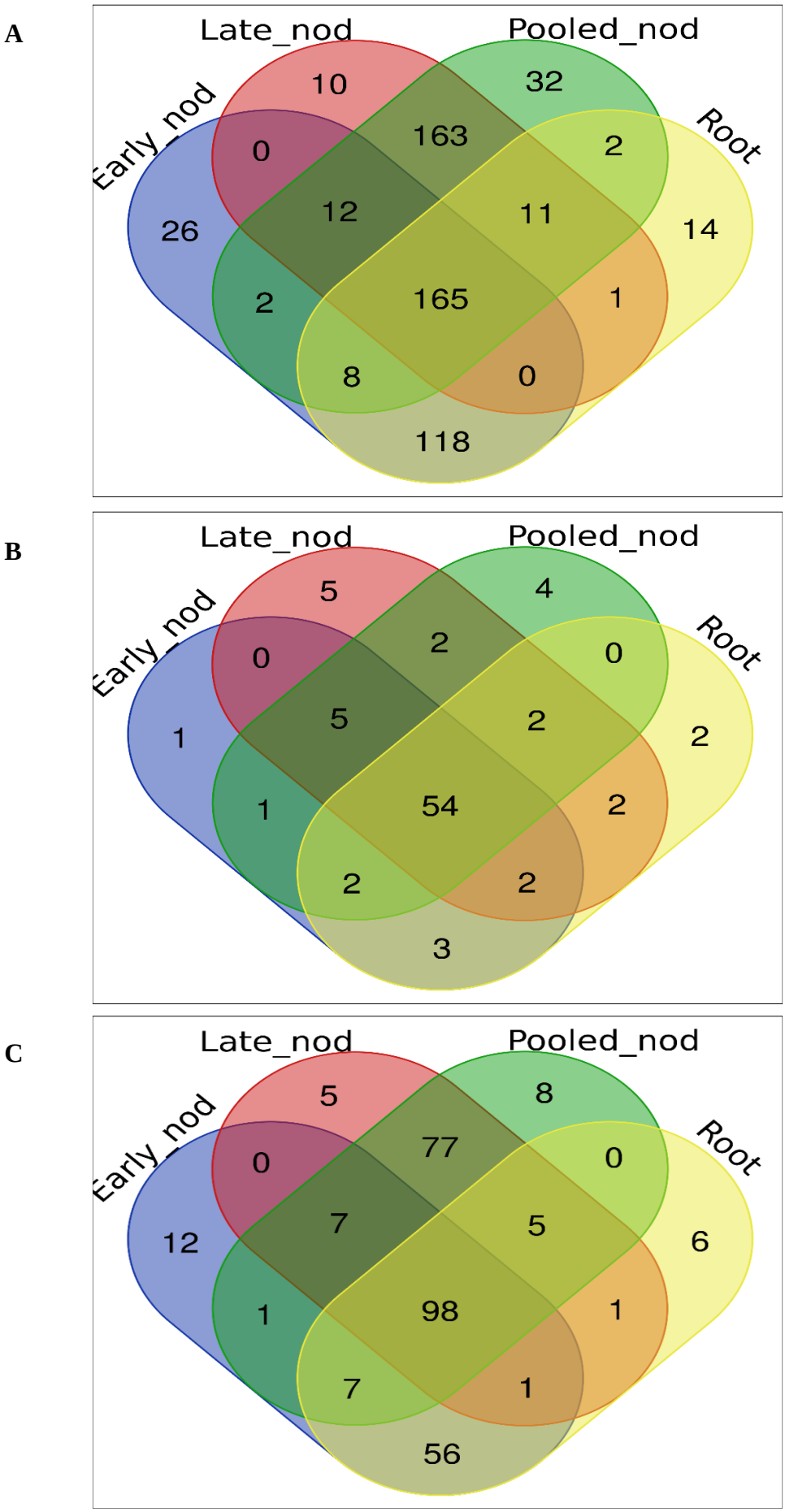


Fig. 7. Venn diagram representation of the cleaved targets, the miRNAs targeting and the associated KEGG pathways. The targets cleaved by miRNAs in the 4 PARE libraries. **(A)** The miRNAs targeting the transcripts identified through PARE libraries (**B**) The differentiation of the KEGG pathways associated with the target transcripts **(C)**

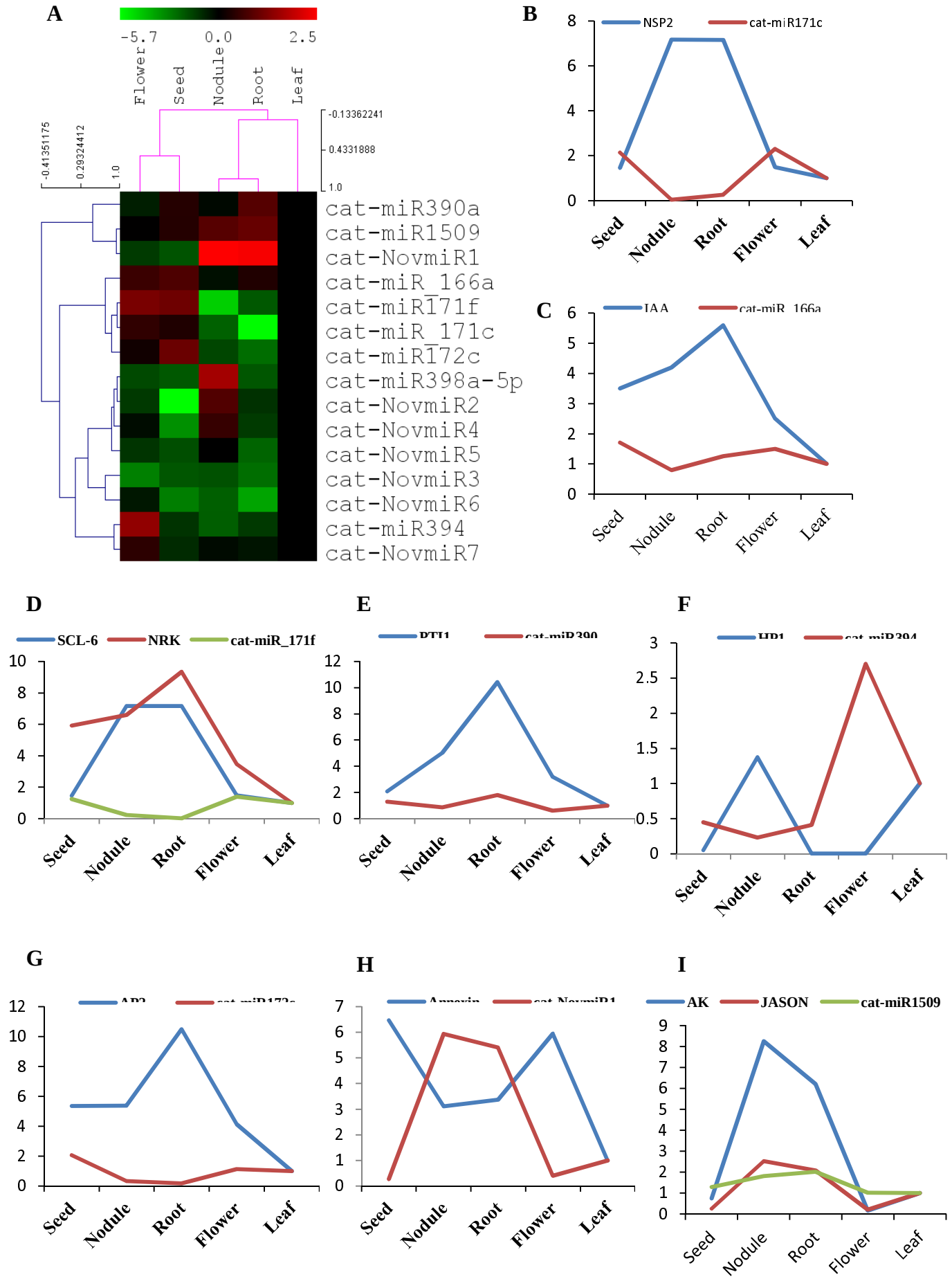
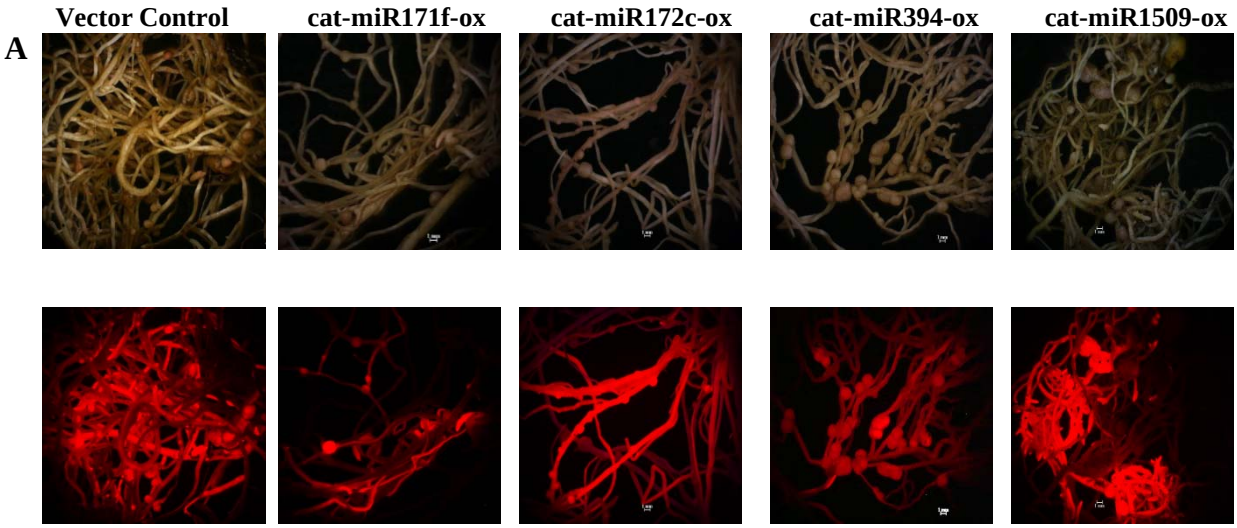


Fig. 8. qRT-PCR based expression profiling. Heat map represents the expression profiling of cat-miRs in 5 different tissues. **(A)** qRT-PCR based expression analysis of miRNA-Target pair in 5 different tissues of chickpea. (**B-I**)



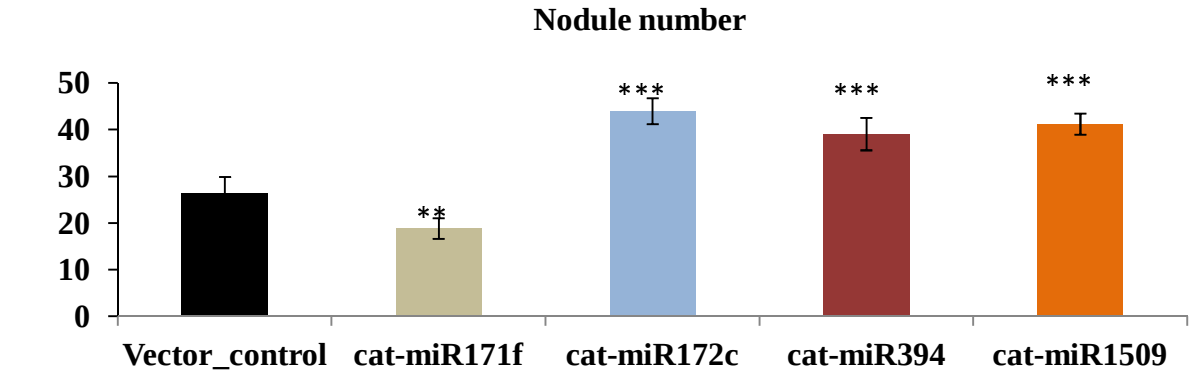


Fig. 9 Ecotopic expression of cat-miRs in chickpea hairy roots resulted in significant changes of nodule numbers. dsRED-expressing root nodules in transgenic roots transformed with vector control and respective miRNAs. (**A**) Relative nodule number formed on chickpea hairy roots expressing cat-miR171f, cat-miR172c, cat-miR394 and cat-miR1509 as compared to roots transformed with the empty vector control. (**B**) ** indicate that the difference is statistically significant at $P < 0.05$ level and *** indicate that the difference is statistically significant at $P < 0.01$ level. Error bar represents SE.

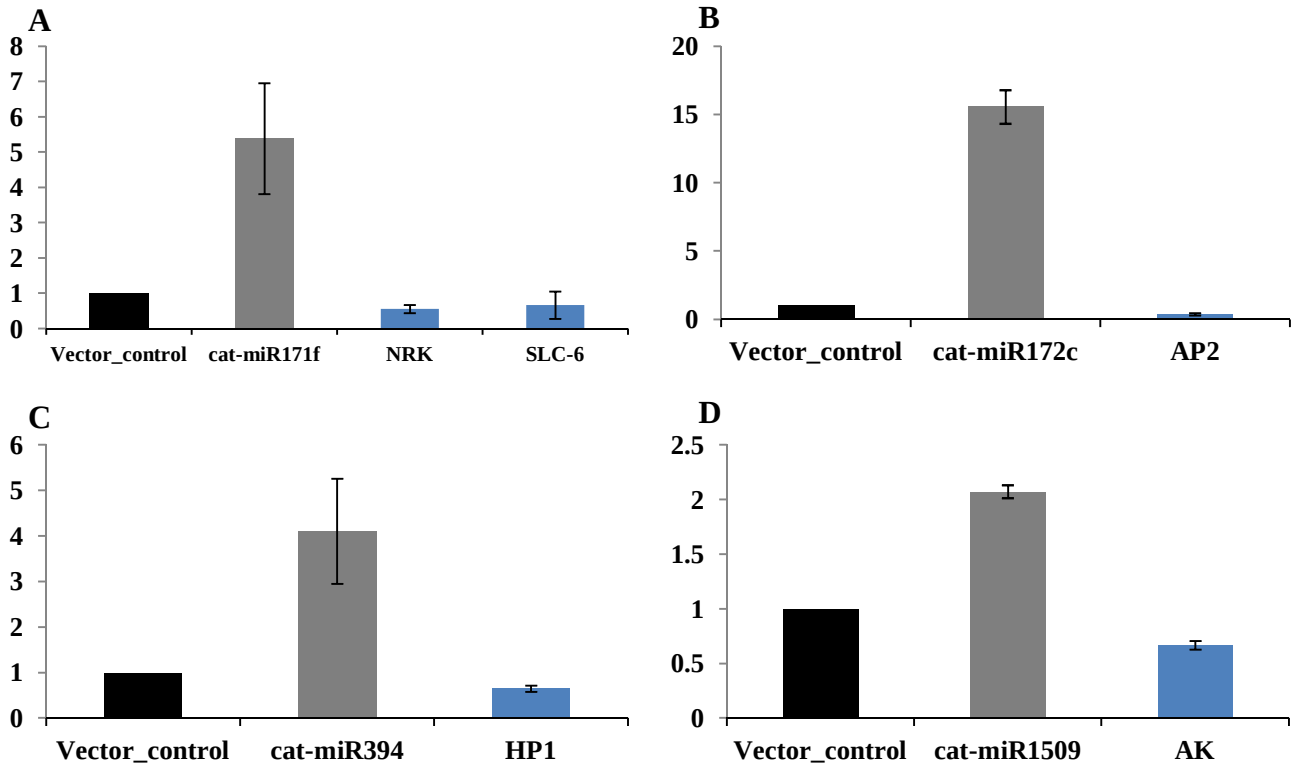


Fig. 10. qRT-PCR based expression of miRNAs and the corresponding target genes in miRNA overexpression lines. qRT-PCR analysis of cat-miR171f and its two targets (SLC-6 and NRK) in cat-miR171f overexpression lines. (**A**) qRT-PCR analysis of cat-miR172c and its target AP2 in cat-miR172c overexpression lines. **(B)** qRT-PCR analysis of cat-miR394 and its target HP1 in cat-miR394 overexpression lines. **(C)** qRT-PCR analysis of cat-miR1509 and its target AP2 in cat-miR1509 overexpression lines. **(D)** Three biological replicates were used. Error bar represents SE. Black bar represent vector control, grey bar represent the miRNA and blue bar represents the corresponding targets. The expression in vector control is indicated by black coloured bars. The grey coloured bars indicates the expression of chickpea miRs in respective overexpression lines and the blue colour bars represent the expression levels of corresponding targets.