



Research article

In silico identification of conserved miRNAs in the genome of fibre biogenesis crop *Corchorus capsularis*

Milad Ahmed^{a,b,1}, Foeaz Ahmed^{a,c,1}, Jamil Ahmed^{a,d,1}, Mst Rubaiat Nazneen Akhand^{a,d}, Kazi Faizul Azim^{a,e}, Md. Abdus Shukur Imran^{a,f}, Syeda Farjana Hoque^{a,f}, Mahmudul Hasan^{a,f,*}^a Faculty of Biotechnology and Genetic Engineering, Sylhet Agricultural University, Sylhet, 3100, Bangladesh^b Department of Animal and Fish Biotechnology, Sylhet Agricultural University, Sylhet, 3100, Bangladesh^c Department of Molecular Biology and Genetic Engineering, Sylhet Agricultural University, Sylhet, 3100, Bangladesh^d Department of Biochemistry and Chemistry, Sylhet Agricultural University, Sylhet, 3100, Bangladesh^e Department of Microbial Biotechnology, Sylhet Agricultural University, Sylhet, 3100, Bangladesh^f Department of Pharmaceuticals and Industrial Biotechnology, Sylhet Agricultural University, Sylhet, 3100, Bangladesh

ARTICLE INFO

Keywords:

Corchorus capsularis

Jute

MicroRNA

ESTs

Natural fibre

Lignin

ABSTRACT

Corchorus capsularis, commonly known as jute occupies the leading position in the production of natural fibre alongside lower environmental threat. Small noncoding ~21 to 24 nucleotides long microRNAs play significant roles in regulating the gene expression as well as different functions in cellular growth and development. Here, the study adopted a comprehensive *in silico* approach to identify and characterize the conserved miRNAs in the genome of *C. capsularis* including functional annotation of specific gene targets. Expressed Sequence Tags (ESTs) based homology search of 3350 known miRNAs of dicotyledons were allowed against 763 non-redundant ESTs of jute genome, resulted in the prediction of 5 potential miRNA candidates belonging five different miRNA families (miR1536, miR9567-3p, miR4391, miR11300, and miR8689). The putative miRNAs were composed of 18 nucleotides having a range of -0.49 to -1.56 MFEI values and 55%–61% of (A + U) content in their pre-miRNAs. A total of 1052 gene targets of putative miRNAs were identified and their functions were extensively analyzed. Most of the gene targets were involved in plant growth, cell cycle regulation, organelle synthesis, developmental process and environmental responses. Five gene targets, namely, NAC Domain Containing Protein, WRKY DNA binding protein, 3-dehydroquinate synthase, S-adenosyl-L-Met-dependent methyl transferase and Vascular-related NAC-Domain were found to be involved in the lignin biosynthesis, phenylpropanoid pathways and secondary wall formation. The present study might accelerate the more miRNA discovery, strengthening the complete understanding of miRNAs association in the cellular basis of lignin biosynthesis towards the production of high standard jute products.

1. Introduction

Jute is the blessing of future economy on account of raising awareness with environmental and SDG (Sustainable Development Goal) issues around all over the world [1]. It is now placed in the second and only after cotton in case of economic importance among the natural fibers and over 95% of the global jute production is served by India, Bangladesh, China, Nepal and Thailand [2]. Jute products are more environmentally sound and their disposal stage is also less harmful to environment than its competing synthetic materials [3]. Moreover, the natural fibre products are likely to be the favorable for combating against global warming [4].

Moreover, on an average farm value of ~US\$2.3 billion is being generated by the jute products annually [5]. Nowadays, scientists are more likely to be interested in developing cost-effective and eco-friendly polymeric compounds from natural resources as implementation of new laws and raising pressure from environmental activists [6, 7]. Recently, genome sequencing of different bast (phloem) fibres plants such as jute (*Corchorus sp.*) and kenaf (*Hibiscus cannabinus L.*) have been completed focusing on the analysis of vital structural and regulatory genes involved in fibre biogenesis [8, 9].

Jute (*Corchorus sp.*) covers about 80% of worldwide bast fibre production [8]. *C. capsularis* and *Corchorus olitorius* are commercially

* Corresponding author.

E-mail address: mhasan.pib@sau.ac.bd (M. Hasan).¹ Authors contributed equally.

produced, though there are more than hundreds of *Corchorus* plant species in the family of Malvaceae [10]. Lignocellulosic bast fibres are extracted from *C. capsularis* for commercial purposes in the different industrial uses and textile sectors [11]. Usually jute fibre is composed of 61 % cellulose, 15.9 % hemicelluloses and 13.5 % lignin [12], but in some cases it could contain as high as 15–22 % of lignin [13]. The high content of lignin gives a unique xylan-type fibre properties of jute fibres which are distinguished from gelatinous-based other bast fibre types originated from hemp, flax and ramie [14]. Jute fibre lignin seems the polymerization of three hydroxycinnamyl alcohols (monolignols), viz., p-coumaryl alcohol, sinapyl alcohol and coniferyl alcohol [15]. However, this high level of lignin content causes several complexities in the formation of jute fibre and its valued product processing. For examples, photo-yellowing of jute products could be happened due to the α -carbonyl group of the lignin composite which reacts with UV radiation. Besides, high lignin content also cause complexity in jute fibre separation as well as maintaining standard quality [16]. The above mentioned features has made the jute fibres less suitable for making divers valued finer fabrics and other standard products, and it mostly happens in the fibres from *C. capsularis*. So, the development of commercial low-lignin fibres has become a priority needs in jute industries as well as global market [17].

Different studies reported several metabolic pathways as well as genes involved in the key processes of fibre formation [18]. But, there are yet limited studies in the molecular basis and regulatory process of lignin biosynthesis in xylan type bast fibre [19]. However, the availability of genomic data and comparative genome based analysis could allow the studies of gene regulation and molecular interaction of different pathways involved in complex fibre formation process.

MicroRNAs (miRNAs), 18–22 nucleotides long small regulatory RNAs could play a significant role in different cellular functions in plants and animals [20]. Regulating gene expression at the different stages of posttranscriptional levels could be interfered by the miRNAs through the interaction with specific gene targets for initiating mRNA inhibition or cleavage of mRNA translation [21]. Though, there are enormous possibilities to work with miRNA for controlling any specific biological process but still limited number of miRNAs has been detected in plants by experimental and computational strategies [22]. As miRNAs are abundantly linked up through evolutionary lineage in both plant and animal species that is why the nature of evolutionary conservation in the miRNAs allows an effective approach to the identification and characterization of newly conserved miRNAs using comparative genomics approach [23].

Though the plethora of economic importance in the view of trading and commercial purposes, there has been no report on miRNAs in *C. capsularis*. In this study, we employed an EST based homology search to identify miRNAs in jute using currently available 826 expressed sequence tags (ESTs) from the NCBI Genbank database. To avoid the difficulties of searching from genomic sequence surveys or unavailability of whole genome sequences, EST based miRNA identification by using specialized software has already been accepted successfully over the other approaches [24]. The present study identified conserved microRNAs in the genome of fibre biogenesis crop jute by using EST based comparative genomics approach.

2. Methods

2.1. Acquisition of EST sequences and reference miRNAs

The genome based BLAST search of *C. capsularis* (GenBank Assembly Accession: GCA_001974805) suggested available ESTs of NCBI database. The suggested ESTs of *C. capsularis* were retrieved from the EST database of NCBI (<http://www.ncbi.nlm.nih.gov/>). Moreover, to search potential miRNAs in *C. capsularis*, all previously known mature miRNA sequences of plant dicotyledons were extracted from miRBase database (<http://www.mirbase.org/>). These miRNAs were employed as the reference

set with the retrieved ESTs of *C. capsularis* genome for identifying conserved miRNAs family in jute. To avoid redundant or overlapping of miRNAs and ESTs, the repeated sequences of miRNAs and ESTs were removed by CD-HIT (<http://weizhongli-lab.org/cd-hit/>) and only the single ones were retained for further study.

2.2. Search for potential miRNAs in the jute genome

The reference miRNAs retrieved from miRBase were used as the query templates for homology search against the ESTs of *C. capsularis* by using BLASTn of NCBI database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) considering all default parameters. Only the top EST hit for each miRNA search was taken. Again, all of ESTs from BLASTn search were screened by CD-HIT for redundancy check.

2.3. Screening for the non-coding miRNA candidates

As the miRNAs are non-protein coding genes, the precursor miRNAs should also be non-protein coding. After redundancy check, the top hits from BLASTn were subjected to BLASTx (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) analysis for removing the protein-coding sequence. A number of non-protein coding EST sequences were screened. By investigating the miRNA precursor determinant properties, these miRNA candidates were allowed for proper testing to finalize the pre-miRNAs.

2.4. Identification of putative miRNAs with reported criteria

The different criteria were used to screen precursor miRNAs based on statistical comparative genomics approaches reported by Zhang et al., 2006 [25], which had already been adopted in many ESTs targeted miRNAs identification and characterization studies [26]. Here, only candidate ESTs adopted the following properties were prioritized as considerable miRNAs in *C. capsularis*: (1) no more than 3 mismatches were permitted in the putative miRNAs with all formerly reported miRNAs of plant species; (2) the length of finalized putative mature miRNAs must be kept in the range of 18 nucleotides; (3) pre-miRNA sequence must fold into an accurate hairpin stem–loop secondary structure and the identified miRNAs should be placed in one arm of the corresponding hairpin stem–loop; (4) the mature miRNAs should allow less than 3 mismatches with the opposite miRNA strand in the other arm; (5) the percentage of (A + U) content of pre-miRNA should be kept within the range of 30%–70%; and (6) the secondary structure of mature miRNAs must have lower minimal (highly negative) folding free energy (MFE) and minimal free energy index (MFEI) value.

2.5. Prediction of pre-miRNA and hair-loop secondary structure of mature miRNAs

The ESTs for putative miRNA prediction were subjected to mirEval (<http://mimirna.centenary.org.au/mireval/>) for determining the pre-miRNA candidates. The minimal length of the pre-miRNA was fixed as 85 nucleotide and these precursor sequences of effective miRNA homologs were tested for secondary structures prediction using the Zuker folding algorithm in the Mfold web server (<http://unafold.rna.albany.edu/?q=mfold>). As the predicted secondary structure should have the higher minimal negative value of minimal folding free energy index (MFEI) and minimal folding free energy (MFE), only the best fitted pre-miRNA candidates were obtained considering all parameters [23, 26, 51]. The MFEI was calculated using the following equation;

Adjusted MFE (AMFE) represented the MFE of 100 nucleotides.

$$\text{AMFE} = (\text{MFE} \div \text{length of RNA sequence}) * 100$$

$$\text{MFEI} = \text{AMFE} / (\text{G} + \text{C}) \%$$

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

2.6. Nomenclature and family annotation of predicted microRNAs

The putative miRNAs were used as query sequence for local blast in miRBase to reconfirm the family of the putative mature miRNAs. The predicted microRNAs were named by a nomenclature followed by miR-Base (<http://www.mirbase.org/>). Sequence alignment of homologous members of predicted miRNA family was investigated in the miRBase search output and rechecked by Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). The percentage of each nucleotide base of putative mature miRNAs and their respective homologue miRNAs from miRBase were also calculated and analyzed.

2.7. Prediction and functional analysis of putative miRNA targets

Due to the unavailability of *C. capsularis* cDNA library for miRNA target prediction, *Arabidopsis thaliana* was considered as the reference organism for prediction and functional annotation of the putative newly detected mature miRNAs in jute genome. The widely used miRNA target prediction tool psRNATarget database (<http://plantgrn.noble.org/psRNATarget/>) was used here, and the mature miRNAs were employed as query against the *A. thaliana* considering the following parameters: (1) maximum expectation value should be 5; (2) complementarity scoring (HSP) size should be kept within the length of 18; (2) target sites' multiplicity must be allowed to 2; (3) range of central mismatch for translational inhibition should be in the range of 10–11 nucleotide; (4) highest level of mismatches at the complementary site 4 except any gaps. The psRNATarget ensures reverse complementary matching between targets transcripts and respective miRNAs, leading to the target site accessibility by calculating unpaired energy (UPE) needed for beginning the secondary structure around the miRNA target site [27]. The schematic outline of the total methodology was illustrated in Figure 1.

3. Results

3.1. Search of potential miRNAs in genome of *C. capsularis*

To identify and characterize the conserved microRNA of *C. capsularis*, a comprehensive EST based study was employed. There were different *in silico* steps where a huge volume of preliminary ESTs of *C. capsularis* were screened to finalize the putative miRNAs with specific gene targets involved in the biological process of jute alongside biogenesis of fibre. The summary of details process with the numerical data was presented in Table 1. The previously known mature miRNAs under the dicotyledons, belonging to twenty plants families (Amaranthaceae, Araliaceae, Asteraceae, Cucurbitaceae, Euphorbiaceae, Fabaceae, Caricaceae, Lamiales, Linaceae, Malvaceae, Myrtaceae, Paeoniaceae, Brassicaceae, Ranunculaceae, Rhizophoraceae, Rosaceae, Rutaceae, Salicaceae, Solanaceae, Vitaceae) were extracted from the microRNA database. In this approach, a total of 6386 of plant miRNA (Supplementary File 1) were retrieved. About 826 ESTs of *C. capsularis* (Supplementary File 2) were retrieved from the ESTs databases of NCBI. Redundant miRNAs and ESTs were removed, and only the unique ones were retained. About 3036 redundant miRNA sequences were removed, retaining 3350 miRNAs containing no repeated sequences were kept for the further investigation (Supplementary File 3). On the other hand, 763 non-redundant ESTs of jute were kept from 826 ESTs (Supplementary File 4). Here, 763 non-redundant jute ESTs were searched against 3350 non-redundant mature miRNA sequences of all dicotyledons. After the removal of redundancy by CDHIT, around 621 ESTs were selected as potential miRNA sequences for *C. capsularis* (Supplementary File 5). Moreover, it had been found that around 36 ESTs were non-coding which could be the potential candidates of miRNA homologs (Supplementary File 6).

3.2. Prediction of pre-miRNAs and hair-loop secondary structure of mature miRNAs

About the 36 non-coding ESTs were investigated to identify the potential miRNA precursors. The different preset criteria were strictly followed in each steps to screen out the putative miRNAs in *C. capsularis*. Here, five ESTs were considered as pre-miRNA by careful evaluation (Table 2 and Supplementary Table 1), while the other 31 ESTs (Supplementary File 7) were filtered out as they failed to fulfill the criteria mentioned earlier. The final five precursor sequences of putative miRNAs were tested for confirming their ability to form secondary structures using the Zuker folding algorithm of MFOLD software (Figure 2). It was observed that all five pre-miRNAs were capable of folding themselves into an appropriate hairpin structure as predicted by mirEval. There was also minimal loop involvement in the hairpin structures of the miRNAs. The four mature miRNAs were incorporated in only one loop of their respective hairpin structure while the other one was in a single-arm without involving any loop (Figure 2). The minimal folding free energy (MFE, dG in kcal/mol) of all the pre-miRNAs were ranged between -59.60 and -14.90 while the percentage of A + U contents of the precursors were ranged between 55% and 65% (Table 2).

3.3. Nomenclature and family annotation of predicted microRNAs

The miRBase database was used to search out and address the microRNAs family of the predicted five putative miRNAs. It had been found that the five miRNAs of jute belong to five different miRNA families (miR1536, miR9567-3p, miR4391, miR11300 and miR8689) and the nomenclature of the predicted miRNAs were done as cca-miR1536, cca-miR9567-3p, cca-miR4391, cca-miR11300, and cca-miR8689. Sequence alignment of putative miRNAs of jute genome and its homologous member from miRBase were visualized in Figure 3. The average each nucleotide base percentage of cca-miR1536, cca-miR9567-3p, cca-miR4391, cca-miR11300, and cca-miR8689 were calculated, and compared with the each nucleotide base percentage of gma-miR1536, bra-miR9567-3p, gma-miR4391, fve-miR11300 and gra-miR8689 respectively. It had been investigated that the percentage of adenine, uracil, guanine and cytosine of both putative mature miRNA of jute and their correspondence reference miRNA were almost in the similar quantity. In addition, the adenine percentage was found higher in the miRNA family of miR9567-3p and miR4391 for whereas other two families, namely, miR1536 and miR8689 showed higher percentages in case of uracil and guanine respectively (Figure 4).

3.4. Prediction and functional analysis of putative miRNA targets

The five novel jute miRNAs were found to be involved with a total of 1052 targets as predicted by the psRNATarget (Supplementary Table 2). Among the around 1052 gene targets, 147 was found to be uncharacterized. All of the gene targets were cross checked where four gene targets were found as the common target for the five putative miRNAs, and those targets were also highly abundant in the target pool ('Leucine-rich repeat family protein; 13 times', 'Transposable element gene; 87 times', 'Protein kinase superfamily protein; 17 times', 'Pentatricopeptide repeat superfamily protein; 9 times'). Comprehensive literature studies were employed to find out the involvement of predicted gene targets in different biological process. There were few gene targets which were found to be directly or indirectly involved with the lignin biosynthesis and secondary wall biogenesis (Table 3) [28–44]. The functions of the common gene targets were also described in Table 3. Moreover, there were three gene targets (alpha/beta-Hydrolases superfamily protein, RING/U-box superfamily protein and P-loop containing nucleoside triphosphate hydrolases superfamily protein coding genes) were found to be involved with four putative miRNAs of jute (Supplementary Table 2).

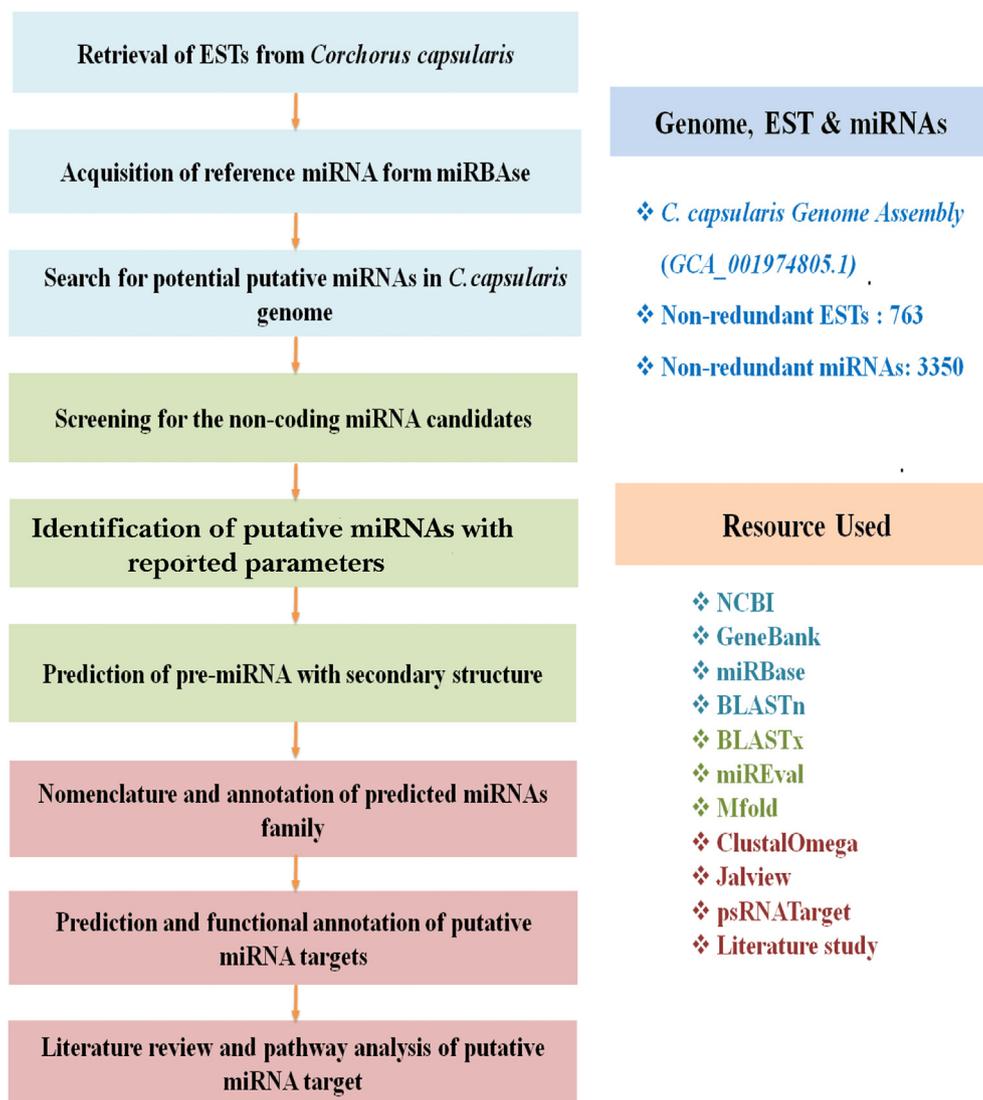


Figure 1. The schematic outline of the total methodology.

Table 1. Steps involved in the characterization of putative miRNA from ESTs of jute genome.

No	Steps	No of ESTs
01	Retrieval of previously known mature miRNA sequences from miRBase	6386
02	Retrieval of EST sequences from jute genome	826
03	Non-redundant miRNA sequences	3350
04	Non-redundant ESTs of jute genome	763
05	Search for potential miRNAs in jute genome by BLASTn	621
06	Candidate ESTs with non-protein coding sequences after Blastx	36
07	Putative pre-miRNA sequences with Hair-loop secondary structure	05
08	Identification of newly identified miRNA family	05
09	Prediction and functional of putative miRNA targets	05
09	Final candidate of conserved newly identified miRNA in the jute genome	05

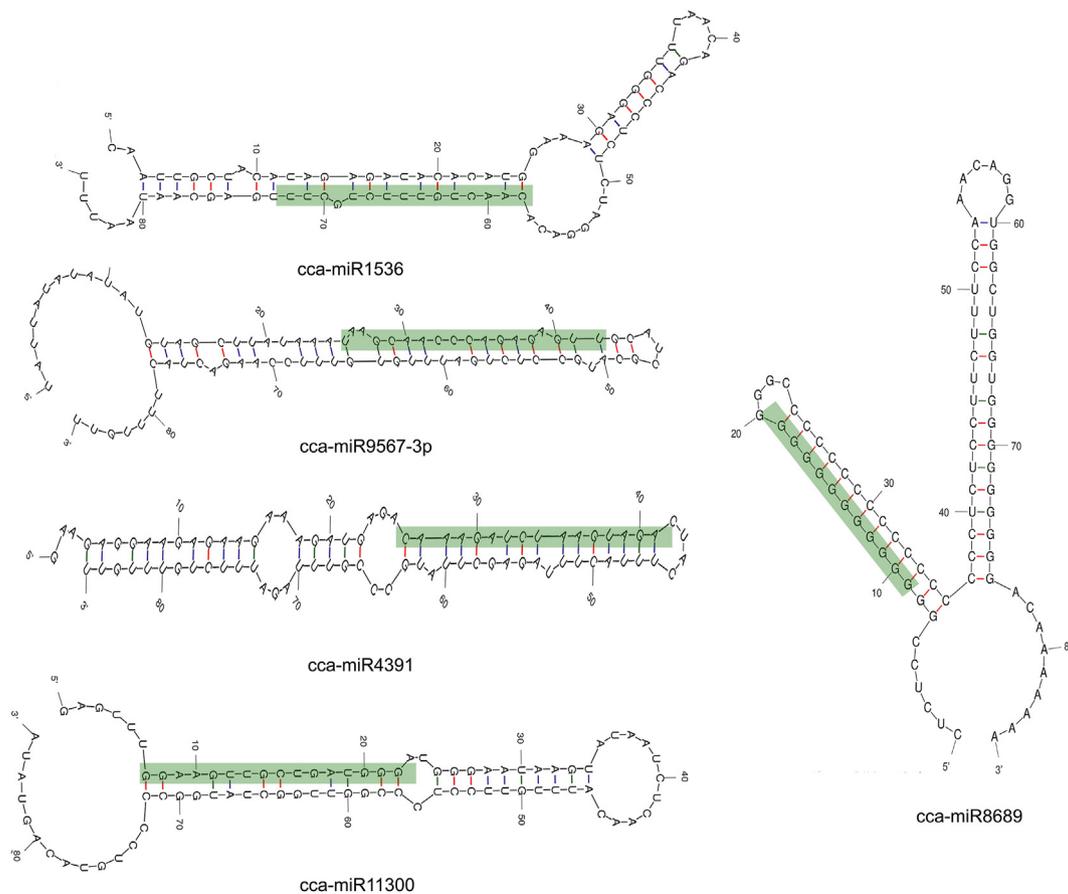
4. Discussion

Identification of conserved microRNA in the genome of *C. capsularis* may have significant roles in the improvement of fibres quality including its mass production. Though xylan type bast fibres are traditionally used in the low quality household products because of its lower smoothness and higher percentage of lignin compounds, but now jute products are being emphasized for the sustainable application for its

eco-friendly nature [45, 46]. Hence, lignin biogenesis pathways and other molecular basis of fibre formation are now getting more importance for its commercial and economic perspective [47]. The regulation of lignin biosynthesis leading to fibre formation could initiate the lower percentage of lignin compound in the jute fibres which is the prerequisite for high quality fibre products. MicroRNAs are now being prioritized in the molecular research for their diverse cellular applications from the post-transcriptional regulation to gene expression [48]. Though plant

Table 2. Putative miRNA in the jute genome with suggested criteria.

Predicted miRNAs of Jute	Accession No of EST	miRNAs of miRBase	Length	Mismatch	A + U content (%)	MFE	MFEI
ACACAACUGUUCUGCUU	GR463689	gma-miR1536	18	3	61	-23.30	-0.58
UAAGCAACCCAGAGAGUU	HS411997	bra-miR9567-3p	18	3	65	-17.60	-0.60
GACAAAGAUCAAGUAGA	FK826577	gma-miR4391	18	3	64	-14.90	-0.49
AGGAAGAGAAGAAAGAUG	FK826577	fve-miR11300	18	3	64	-22.00	-0.73
AGUUGCUGAUGGGAUGGG	GH985240	gra-miR8689	18	3	55	-59.60	-1.56

**Figure 2.** Secondary structures of final putative miRNA by using the Zuker folding algorithm of MFOLD software.

originated miRNAs are now being prioritized in the scientific studies [49], there are minimal approaches in the identification and functional characterization of miRNAs in *C. capsularis*. The present study was carried out to identify the putative conserved microRNA candidates from the genome of *C. capsularis* by employing different screening process. Plant miRNAs are found to be conserved in the different evolutionary convergent species [50]. In the present study, previously reported miRNAs of dicotyledons were prioritized to conduct homology based search for identifying the putative miRNA homologs of *C. capsularis* in the available ESTs of jute genome. Though, few previous miRNA studies of higher plants used only the miRNAs of *Arabidopsis thaliana* for their homology search [51], but present study considered around 3350 non-redundant mature miRNAs of dicotyledons plants for enhancing the strength of broad spectrum homology search for conserved miRNAs in the jute genome.

Retrieved mature miRNAs of dicotyledons were allowed to homology search with 763 non-redundant ESTs of *C. capsularis* where 621 ESTs were identified for the further screening process. Non-coding ESTs were prioritized for being the putative miRNAs candidates [52] and only 36 non-coding ESTs were finalized in the present study. Different criterias

were set up for screening the ESTs in order to identify the putative miRNAs of jute [53]. Though it had been reported that 10,000 ESTs in plants could suggest single miRNA in plant species [54], five miRNAs have been characterized in *C. capsularis* genome through scrutinizing a series of screening processes from initial 826 ESTs. The putative miRNAs were found with 18 nucleotide length which are also observed in the other studies [55]. The pri-miRNAs which are the precursors of pre-miRNAs contain short open reading frame sequences that encode miRNA-encoded peptide (miPEP) that enhance the accumulation of their corresponding mature miRNAs [56]. The pre-miRNA were checked for the ability of folding into an accurate hairpin secondary structure alongside with the perfect positioning of mature miRNAs in the arm of hairpin structure (Figure 2). Minimal free energy index (MFEI) indicates the probable miRNA (average MFEI = -0.65) candidates with respect to other RNAs such as tRNAs (MFEI = -0.64) and rRNAs (MFEI = -0.59) [57]. The MFEI value of the newly identified five putative miRNAs of jute were found within the average range (-0.49, -0.58, -0.6, -0.73, and -1.56). Here, (A + U) content of pre-miRNA were found between the range of 55%–61%, which was also observed in the other higher plants miRNAs [58]. Five conserved miRNA families (miR1536, miR9567-3p, miR4391,

cca-miR1536	1	ACACAACUGUUUCUGCUU	18
gma-miR1536	18	ACACAUUUGUCUCUGCUU	1
CCA-miR9567-3P	1	UAAGCAACCCAGAGAGUU	18
bra-mi9567-3P	19	UAAGCAACACAUAUAGUU	2
cca-miR4391	1	GACAAAGAUCUAAGUAGA	18
gma-miR4391	5	GGCAAAGAACUAAGAAGA	22
cca-MIR11300	1	AGGAAGAGAAGAAAGAUG	18
fve-MIR11300	21	AGGAAGAGAACAGUGAUG	4
cca-miR8689	1	AGUUGCUGAUGGGAUGGG	18
gra-miR8689	3	AGGUGCUGAUGGCGUGGG	20

Figure 3. Alignments of the putative miRNAs of jute genome and its homologues from respective microRNA family.

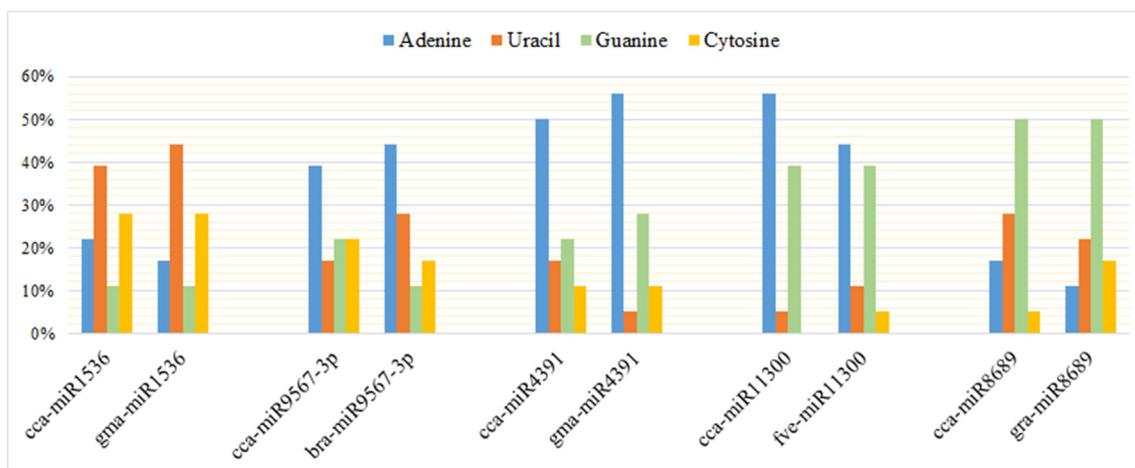


Figure 4. Overall nucleotide compositions (%) of putative miRNA from *Corchorus capsularis* compared with the homologues mRNA of miRBase.

miR11300 and miR8689) were found in the jute genome which were also reported in the different plant families [59] and the predictive putative miRNAs were entitled as cca-miR1536, cca-miR9567-3p, cca-miR4391, cca-miR11300, and cca-miR8689. Multiple sequence alignment and comparative nucleotide composition studies exhibited that the putative miRNAs of jute were highly conserved in the other organisms which supported the evolutionary convergent relationships among the higher plant species [60].

In addition, a total of 1052 genes were identified which could be inhibited by the predicted mature miRNAs of jute. Among this huge gene pool, there were five genes which were somehow involved in the secondary cell wall formation and lignin biosynthesis process. ‘NAC Domain

Containing Protein’, ‘WRKY DNA binding protein’, ‘3-dehydroquinase synthase’, ‘S-adenosyl-L-Met-dependent methyl transferase’ and ‘Vascular-related NAC-Domain’ coding genes associated with the putative miRNA of jute were studied exclusively for their important roles in the lignin biosynthesis or secondary cell wall formation. The study revealed that cca-miR4391 and cca-miR11300 might inhibit the ‘NAC Domain Containing Protein’ whereas cca-miR1536, cca-miR4391 and cca-miR8689 might cleave the ‘S-adenosyl-L-Met-dependent methyl transferase’. Secondary cell wall formation associated NAC protein and its functional homologues are crucial regulatory points that turn on a subset of transcription factors which also directly activate the expression of SCW biosynthetic genes [28, 29]. The diverse variety of

Table 3. Putative miRNA targets in the crucial cellular process of *Corchorus capsularis*.

Target protein	Putative Jute miRNA	Target Acc. No	Inhibition Type	Biological Functions
NAC Domain Containing Protein	cca-miR4391 cca-miR11300	AT2G24430 AT5G64060	Cleavage	Lignin biosynthesis process [28, 29] Secondary cell wall formation [28, 29]
S-adenosyl-L-Met-dependent methyl transferase	cca-miR1536, cca-miR4391, cca-miR8689	AT3G44870 AT1G04430 AT4G01850 AT1G69523 AT1G69526	Cleavage	Lignin biosynthesis pathways [30] Formation of plant secondary cell wall, pigments, isoflavonoid, phytoestrogens and chalcone modulation factors [31]
WRKY DNA binding protein	cca-miR9567, cca-miR8689	AT1G13960 AT5G46350 AT2G47260	Cleavage	Regulation of phenylpropanoid pathway involved in lignin biosynthesis [32] Biosynthesis of important pharmaceutical, aromatherapy and biofuel [33] Production of phytohormones and phytoalexins [34] Determination of Seed size and colour [35]
3-dehydroquinase synthase	cca-miR1536	AT3G28760	Cleavage	Lignin biosynthesis pathways [36] Shikimate and aromatic amino acid biosynthetic pathways [36]
Vascular-related NAC-Domain (VND)-4	cca-miR9567-3p	AT1G12260	Cleavage	Upregulation of the genes required for secondary cell wall deposition [37] Regulation of cell death and prevention of secondary wall defects in the fibers [38]
Leucine-rich repeat (LRR) family protein	cca-miR1536, cca-miR9567-3p, cca-miR4391, cca-miR4391, cca-miR8689	AT1G72460 AT4G29240 AT5G22320 AT1G49750 AT1G53350	Cleavage	Ligand recognition and involved in the protein-protein interaction process [39]
Transposable element gene	cca-miR1536, cca-miR9567-3p, cca-miR4391, cca-miR4391, cca-miR8689	AT4G04550 AT1G44045 AT1G35465 AT3G48235 AT3G44705	Cleavage	Cell cycle regulation, modulating vesicle transport, regulating cellular metabolism and maintenance of energy homeostasis process [40]
Protein kinase superfamily protein	cca-miR1536, cca-miR9567-3p, cca-miR4391, cca-miR4391, cca-miR8689	AT2G42960 AT1G18390 AT5G56890 AT4G24480 AT4G00955	Cleavage	Components of different signalling networks [41] Cell cycle regulation [42] Developmental processes, modulating vesicle transport and in regulating cellular metabolism [43]
Pentatricopeptide repeat (PPR) superfamily protein	cca-miR1536, cca-miR9567-3p, cca-miR4391, cca-miR4391, cca-miR8689	AT3G54980 AT5G61400 AT3G08820 AT4G14850 AT1G11710	Cleavage	Organelle biogenesis, photosynthesis, respiration, plant development and environmental responses [44]

'S-adenosyl-L-Met-dependent methyl transferase' has been observed in the plant. During the biogenesis of deferent plant secondary metabolites alongside lignin molecules, 'S-adenosyl-L-Met-dependent methyl transferase' usually act on Phe-derived substrates. This enzymes has also been reported for its contribution in the synthesis of phytoestrogens, allelochemicals, flower pigments, antimicrobial compounds and isoflavonoid [30, 31]. 'WRKY DNA binding protein' could be inhibited by the predicted cca-miR9567 and cca-miR8689 of *C. capsularis*, and this protein could regulate the phenylpropanoid pathway and contribute in the biosynthesis of important pharmaceutical, aromatherapy and biofuel [32, 33]. WRKYs are exclusively needed for proper cellular expression of participants genes in the biosynthesis of lignin compounds, and these are also found to be involved in the synthesis of phytoalexins, phytohormones, and other chemicals related in defense mechanism [34, 35]. Again, the putative cca-miR1536 and cca-miR9567 were found to inhibit the '3-dehydroquinase' and 'S-adenosyl-L-Met-dependent methyl transferase' where '3-dehydroquinase' had significance contribution in the shikimate and aromatic amino acid biosynthetic pathways [36]. In addition, Vascular-related NAC-Domain (VND)' could be inhibited by the predicted cca-miR9567-3p. It had also been reported that VND genes could regulate the secondary wall biosynthesis process in *Arabidopsis*, specifically in the xylem conducting cells [37]. The expression of VND4 gene was found mostly in the vessels and it was able to turn on transcription factors involved in secondary wall-associated functions and programmed cell death. VND4 was also found to prevent the defects in the fibre of secondary wall [38]. Though lignin provides mechanical strength to save fibre from biological or chemical degradation, it assumes to easy separation of lignin during fibre process for making quality products [61]. Here, the putative potential miRNAs were found to target essential genes involved in the lignin or secondary wall synthesis alongside with different cellular process. The miRNA targeted regulation over the 63, expression has already been reported in different plant species [62, 63, 64], hence, the putative miRNA of jute might exhibit regulatory roles over the important cellular expression through *in vivo* and *in vitro* validation. EST based miRNA prediction is getting more prioritized in different crops as the regulatory profile of targeted miRNAs could impact on the better crop production as well as significant improvement in the desired characteristics of the economically important crop plants [65, 66, 67, 68]. That is why, the above study could have significant importance in the fibre industry as the contributory roles of putative miRNAs of *C. capsularis* have been identified in the lignin biosynthesis process.

5. Conclusion

The characterized jute miRNA's may be of great relevance in trying to understand how miRNAs are associated with the cellular basis of lignin biosynthesis towards the production of high standard jute products. As the screening and characterization of mature miRNAs have been investigated in *C. capsularis*, it seems the predicted mature miRNAs and target genes might be larger. Therefore, these miRNAs and the corresponding targets are needed to be further studied for complete understanding of miRNA derived regulation over the production of high quality fibres. The current study concluded with the suggestions of putative miRNAs in the jute genome with their correspondent gene targets, and also reflected some of the key regulatory proteins involved in the lignin biosynthesis that might be controlled by miRNA dependent cellular actions. In future, other researchers working on *in silico* identification of microRNAs from dicotyledones family can benefit significantly to accelerate the discovery of conserved miRNAs through *in silico* approaches.

Declarations

Author contribution statement

Mahmudul Hasan: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Milad Ahmed, Foeaz Ahmed, Jamil Ahmed: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Mst Rubaiat Nazneen Akhand, Kazi Faizul Azim, Abdus Shukur Imran: Performed the experiments; Contributed reagents, materials, analysis tools or data.

Syeda Farjana Hoque: Performed the experiments.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

বাংলাদেশের স্বাধীনতার সুবর্ণজয়ন্তীতে আমরা আমাদের এই প্রকাশনাটি পাঠের জন্যে নাম সন্কেয়নেসেরে নতুনত্বদানকারী প্রফসের ড. মাকসুদুল আলম ও তাঁর পুরনো 'স্বপ্নযাত্রা' দলকে উসর্গ করছি।

Supplementary content related to this article has been published online at <https://doi.org/10.1016/j.heliyon.2021.e06705>.

Acknowledgements

Authors would like to acknowledge the Faculty of Biotechnology and Genetic Engineering, Sylhet Agricultural University for the technical support of the project.

References

- [1] A. Uygur, The future of organic fibres, *Eur. J. Sustain. Dev. Res.* 2 (2017) 164–172.
- [2] R. Blackburn, *Biodegradable and Sustainable Fibres*, first ed., Elsevier, USA, 2005.
- [3] J.E. van-Dam, Natural fibres and the environment: environmental benefits of natural fibre production and use, in: *Proceedings of the Symposium on Natural Fibres: Common Fund for Commodities*, 2008, pp. 3–17.
- [4] M.S. Islam, S.K. Ahmed, The impacts of jute on environment: an analytical review of Bangladesh, *J. Environ. Earth Sci.* 5 (2012) 24–31.
- [5] *Statistical Bulletin-2014*, Food and Agriculture Organization, United Nations, 2014.
- [6] V.K. Thakur, M.K. Thakur, R.K. Gupta, Review: raw natural fiber-based polymer composites, *Int. J. Polym. Anal. Char.* 19 (2014) 256–271.
- [7] H. Li, M.M. Sain, High stiffness natural fiber-reinforced hybrid polypropylene composites, *Polym. Plast. Technol. Eng.* 42 (2003) 853–862.
- [8] M.S. Islam, J.A. Saito, E.M. Emdad, B. Ahmed, M.M. Islam, A. Halim, Q.M. Hossen, M.Z. Hossain, R. Ahmed, M.S. Hossain, S.M. Kabir, Comparative genomics of two jute species and insight into fibre biogenesis, *Nat. Plants* 3 (2017) 1–7.
- [9] L. Zhang, Y. Xu, X. Zhang, X. Ma, L. Zhang, Z. Liao, Q. Zhang, X. Wan, Y. Cheng, Zhang, D. Li, The genome of kenaf (*Hibiscus cannabinus* L.) provides insights into bast fibre and leaf shape biogenesis, *Plant Biotechnol. J.* (2020) 1–14.
- [10] T. Furumoto, R. Wang, K. Kokazaki, F. Afm, M.I. Ali, A. Kondo, H. Fukui, Antitumor promoters in leaves of jute (*C. capsularis* and *C. olitorius*), *Food Sci. Technol. Res.* 8 (2002) 239–243.
- [11] D.A. Baker, T.G. Rials, Recent advances in low-cost carbon fiber manufacture from lignin, *J. Appl. Polym. Sci.* 130 (2013) 713–728.
- [12] D. Sarkar, M.K. Sinha, A. Kundu, C.S. Kar, A. Saha, L.L. Kharbikar, B.S. Mahapatra, Why is ramie the strongest yet stiffest of bast fibres? *Curr. Sci.* 98 (2010) 1570–1572.
- [13] G. Sengupta, P. Palit, Characterization of a lignified secondary phloem fibre-deficient mutant of jute (*C. capsularis*), *Ann. Bot.* 93 (2004) 211–220.
- [14] J.H. Meshram, P. Palit, On the role of cell wall lignin in determining the fineness of jute fibre, *Acta Physiol. Plant.* 35 (2013) 1565–1578.
- [15] R. Vanholme, B. Demedts, K. Morreel, J. Ralph, W. Boerjan, Lignin biosynthesis and structure, *Plant Physiol.* 153 (2010) 895–905.
- [16] J.C. Del Rio, J. Rencoret, G. Matques, J. Li, G. Gellerstedt, J. Jiménez-Barbero, A.T. Martínez, A. Gutiérrez, Structural characterization of the lignin from jute (*C. capsularis*) fibers, *J. Agric. Food Chem.* 57 (2009) 10271–10281.

- [17] S.B. Choudhary, I. Chowdhury, R.K. Singh, S.P. Pandey, H.K. Sharma, A.A. Kumar, P.G. Karmakar, N. Kumari, J. Souframanien, S.J. Jambhulkar, Morphological, histobiochemical and molecular characterisation of low lignin phloem fibre (llpf) mutant of dark jute (*Corchorus olitorius* L.), Appl. Biochem. Biotechnol. 183 (2017) 980–992.
- [18] E.W. Taliencio, D. Boykin, Analysis of gene expression in cotton fiber initials, BMC Plant Biol. 7 (2007) 22 (RC).
- [19] P. Chiaiese, G. Ruotolo, A. Di Matteo, A. De Santo Virzo, A. De Marco, E. Filippone, Cloning and expression analysis of kenaf (*Hibiscus cannabinus* L.) major lignin and cellulose biosynthesis gene sequences and polymer quantification during plant development, Ind. Crops Prod. 34 (2011) 1072–1078.
- [20] P.B. Kwak, Q.Q. Wang, X.S. Chen, C.X. Qiu, Z.M. Yang, Enrichment of a set of microRNAs during the cotton fiber development, BMC Genom. 10 (2009) 457.
- [21] K.A. Khade, M. Panigrahi, S.F. Ahmad, A. Chauhan, P. Kumar, B. Bhushan, Cloning and characterization of Bubaline mammary miRNAs: an in silico approach, Mol. Biol. Rep. 46 (2019) 1257–1262.
- [22] F. Xie, D.C. Jones, Q. Wang, R. Sun, B. Zhang, Small RNA sequencing identifies miRNA roles in ovule and fibre development, Plant Biotechnol. J. 13 (2015) 355–369.
- [23] C.X. Qiu, F.L. Xie, Y.Y. Zhu, K. Guo, S.Q. Huang, L. Nie, Z.M. Yang, Computational identification of microRNAs and their targets in *Gossypium hirsutum* expressed sequence tags, Gene 395 (2007) 49–61.
- [24] W. Jike, G. Sablok, G. Bertorelle, M. Li, C. Varotto, In silico identification and characterization of a diverse subset of conserved microRNAs in bioenergy crop *Arundodonax* L, Sci. Rep. 8 (2018) 1–3.
- [25] B.H. Zhang, X.P. Pan, S.B. Cox, G.P. Cobb, T.A. Anderson, Evidence that miRNAs are different from other RNAs, Cell. Mol. Life Sci. 63 (2006) 246–254.
- [26] X.J. Wang, J.L. Reyes, N.H. Chua, T. Gaasterland, Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets, Genome Biol. 5 (2004) 65.
- [27] X. Dai, P.X. Zhao, psRNATarget: a plant small RNA target analysis server, Nucleic Acids Res. 39 (2011) 155–159.
- [28] M. Kubo, M. Udagawa, N. Nishikubo, G. Horiguchi, M. Yamaguchi, J. Ito, T. Mimura, H. Fukuda, T. Demura, Transcription switches for protoxylem and metaxylem vessel formation, Genes Dev. 19 (2005) 1855–1860.
- [29] Y. Yang, C.G. Yoo, W. Rottmann, K.A. Winkler, C.M. Collins, L.E. Gunter, S.S. Jawdy, X. Yang, Y. Pu, A.J. Ragauskas, G.A. Tuskan, PdWIND3A, a wood-associated NAC domain-containing protein, affects lignin biosynthesis and composition in *Populus*, BMC Plant Biol. 19 (2019) 1–2.
- [30] L. Li, J.L. Popko, X.H. Zhang, K. Osakabe, C.J. Tsai, C.P. Joshi, V.L. Chiang, A novel multifunctional O-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine, Proc. Natl. Acad. Sci. 94 (1997) 5461–5466.
- [31] N.A. Eckardt, Probing the mysteries of lignin biosynthesis: the crystal structure of caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase provides new insights, Plant Cell 14 (2002) 1185–1189.
- [32] H. Wang, U. Avci, J. Nakashima, M.G. Hahn, F. Chen, R.A. Dixon, Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants, Proc. Natl. Acad. Sci. 107 (2010) 22338–22343.
- [33] P. Tripathi, R.C. Rabara, P.J. Rushton, A systems biology perspective on the role of WRKY transcription factors in drought responses in plants, Planta 239 (2014) 255–266.
- [34] X. Liang, X. Chen, C. Li, J. Fan, Z. Guo, Metabolic and transcriptional alternations for defense by interfering OsWRKY62 and OsWRKY76 transcriptions in rice, Sci. Rep. 7 (2017) 1–5.
- [35] S. Besseau, J. Li, E.T. Palva, WRKY54 and WRKY70 co-operate as negative regulators of leaf senescence in *Arabidopsis thaliana*, J. Exp. Bot. 63 (2012) 2667–2679.
- [36] A. Chakraborty, D. Sarkar, P. Satya, P.G. Karmakar, N.K. Singh, Pathways associated with lignin biosynthesis in lignomaniac jute fibres, Mol. Genet. Genom. 290 (2015) 1523–1542.
- [37] M. Yamaguchi, N. Mitsuda, M. Ohtani, M. Ohme-Takagi, K. Kato, T. Demura, VASCULAR-RELATED NAC-DOMAIN 7 directly regulates the expression of a broad range of genes for xylem vessel formation, Plant J. 66 (2011) 579–590.
- [38] J. Zhou, R. Zhong, Z.H. Ye, Arabidopsis NAC domain proteins, VND1 to VND5, are transcriptional regulators of secondary wall biosynthesis in vessels, PLoS One 9 (2014) 105726.
- [39] D.A. Cushing, N.R. Forsthoefel, D.R. Gestaut, D.M. Vernon, Arabidopsis emb175 and other ppr knockout mutants reveal essential roles for pentatricopeptide repeat (PPR) proteins in plant embryogenesis, Planta 221 (2005) 424–436.
- [40] M. Sahebi, M.M. Hanafi, A.J. van Wijnen, D. Rice, M.Y. Rafii, P. Azizi, M. Osman, S. Taheri, M.F. Bakar, M.N. Isa, Y.M. Noor, Contribution of transposable elements in the plant's genome, Gene 665 (2018) 155–166.
- [41] J.M. Stone, J.C. Walker, Plant protein kinase families and signal transduction, Plant Physiol. 108 (1995) 451–457.
- [42] N.G. Halford, S.J. Hey, Snf1-related protein kinases (SnRKs) act within an intricate network that links metabolic and stress signalling in plants, Biochem. J. 419 (2009) 247–259.
- [43] M.D. Lehti-Shiu, S.H. Shiu, Diversity, classification and function of the plant protein kinase superfamily, Philos. Trans. R. Soc. Lond. B Biol. Sci. 367 (2012) 2619–2639.
- [44] A. Barkan, I. Small, Pentatricopeptide repeat proteins in plants, Annu. Rev. Plant Biol. 65 (2014) 415–442.
- [45] D.S. Bajwa, G. Pourhashem, A.H. Ullah, S.G. Bajwa, A concise review of current lignin production, applications, products and their environmental impact, Ind. Crop. Prod. 139 (2019) 111526.
- [46] A. Agrawal, N. Kaushik, S. Biswas, Derivatives and applications of lignin- an insight, SciTech J. 1 (2014) 30–36.
- [47] G. Zhang, Y. Zhang, J. Xu, X. Niu, J. Qi, A. Tao, L. Zhang, P. Fang, L. Lin, J. Su, The CCoAOMT1 gene from jute (*C. capsularis* L.) is involved in lignin biosynthesis in *Arabidopsis thaliana*, Gene 546 (2014) 398–402.
- [48] L. Nian, L. Tu, W. Tang, W. Gao, K. Lindsey, X. Zhang, Small RNA and degradome profiling reveals a role for miRNAs and their targets in the developing fibers of *Gossypium barbadense*, Plant J. 80 (2014) 331–344.
- [49] P. Mohanpuria, N. Duhan, N.K. Sarao, M. Kaur, M. Kaur, In silico identification and validation of potential microRNAs in Kinnow Mandarin (*Citrus reticulata* Blanco), Interdiscip. Sci. 10 (2018) 762–770.
- [50] D.P. Bartel, MicroRNAs: genomics, biogenesis, mechanism, and function, Cell 116 (2004) 281–297.
- [51] A. Pani, R.K. Mahapatra, N. Behera, P.K. Naik, Computational identification of sweet wormwood (*Artemisia annua*) microRNA and their mRNA targets, Genom. Proteom. Bioinf. 9 (2011) 200–210.
- [52] S. Asha, J. Nisha, E.V. Soniya, In Silico characterisation and phylogenetic analysis of two evolutionarily conserved miRNAs (miR166 and miR171) from black pepper (*Piper nigrum* L.), Plant Mol. Biol. Rep. 31 (2013) 707–718.
- [53] B.H. Zhang, X.P. Pan, Q.L. Wang, G.P. Cobb, T.A. Anderson, Identification and characterization of new plant microRNAs using EST analysis, Cell Res. 15 (2005) 336–360.
- [54] N.P. Vishwakarma, V.J. Jadeja, Identification of miRNA encoded by *Jatropha curcas* from EST and GSS, Plant Signal. Behav. 8 (2013) 23152.
- [55] S.M. Karani, A. Bahari, A. Eskandari, D. Koolivand, In Silico Identification of miRNA Encoded by Heterodaglycines Based on Database of EST, Iranian Plant Protection Congress, 2016, p. 233.
- [56] L. Dominique, J. Couzigou, H.S. Clemente, Y. Martinez, C. Dunand, G. Bécard, J. Combier, Primary transcripts of microRNAs encode regulatory peptides, Nature 520 (2015) 90–93.
- [57] B. Zhang, X. Pan, E.J. Stellwag, Identification of soybean microRNAs and their targets, Planta 229 (2008) 161–182.
- [58] M.Y. Barozai, M. Din, I.A. Baloch, Structural and functional based identification of the bean (*Phaseolus*) microRNAs and their targets from expressed sequence tags, J. Struct. Funct. Genomics 14 (2013) 11–18.
- [59] H.J. Kim, K.H. Baek, B.W. Lee, D. Choi, C.G. Hur, In silico identification and characterization of microRNAs and their putative target genes in Solanaceae plants, Genome 54 (2011) 91–98.
- [60] M.J. Axtell, J.L. Bowman, Evolution of plant microRNAs and their targets, Trends Plant Sci. 13 (2008) 343–349.
- [61] L. Li, J.K. Popko, X.H. Zhang, K. Osakabe, C. J Tsai, C.P. Joshi, V.L. Chiang, A novel multifunctional O-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine, Proc. Natl. Acad. Sci. 94 (1997) 5461–5466.
- [62] Z. Ma, C. Coruh, M.J. Axtell, Arabidopsis lyrata small RNAs: transient MIRNA and small interfering RNA loci within the Arabidopsis genus, Plant Cell Rep. 22 (2010) 1090–1103.
- [63] H. Vaucheret, F. Vazquez, P. Crété, D.P. Bartel, The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development, Genes Dev. 18 (2004) 1187–1197.
- [64] B. Zhang, Q. Wang, K. Wang, X. Pan, F. Liu, T. Guo, G.P. Cobb, T.A. Anderson, Identification of cotton microRNAs and their targets, Gene (2007) 26–37.
- [65] G. Ayachit, H. Pandya, J. Das, miRDetect, A combinatorial approach for automated detection of novel miRNA precursors from plant EST data using homology and Random Forest classification, Genomics 112 (2020) 3201–3206.
- [66] B. Hajieghrari, N. Farrokhi, Investigation on the conserved MicroRNA genes in higher plants, Plant Mol. Biol. Rep. 15 (2020) 1–4.
- [67] D. B. Krishnatreya, P.M. Baruah, B. Dowarah, K.S. Bordoloi, H. Agarwal, N. Agarwala, Mining of miRNAs from EST data in *Dendrobium nobile*, Bioinformatics 16 (2020) 245.
- [68] P.G. Patil, N.V. Singh, S. Parashuram, A. Bohra, D.M. Mundewadikar, V.R. Sangure, K.D. Babu, J. Sharma, Genome wide identification, characterization and validation of novel miRNA-based SSR markers in pomegranate (*Punica granatum* L.), Physiol. Mol. Biol. Plants 3 (2020) 1–4.