

Research Article

High-Throughput Sequencing Reveals Diverse Sets of Conserved, Nonconserved, and Species-Specific miRNAs in Jute

**Md. Tariqul Islam, Ahlan Sabah Ferdous, Rifat Ara Najnin,
Suprovath Kumar Sarker, and Haseena Khan**

Molecular Biology Laboratory, Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka 1000, Bangladesh

Correspondence should be addressed to Haseena Khan; haseena@du.ac.bd

Received 11 November 2014; Revised 13 February 2015; Accepted 23 February 2015

Academic Editor: Mohamed Salem

Copyright © 2015 Md. Tariqul Islam et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

MicroRNAs play a pivotal role in regulating a broad range of biological processes, acting by cleaving mRNAs or by translational repression. A group of plant microRNAs are evolutionarily conserved; however, others are expressed in a species-specific manner. Jute is an agroeconomically important fibre crop; nonetheless, no practical information is available for microRNAs in jute to date. In this study, Illumina sequencing revealed a total of 227 known microRNAs and 17 potential novel microRNA candidates in jute, of which 164 belong to 23 conserved families and the remaining 63 belong to 58 nonconserved families. Among a total of 81 identified microRNA families, 116 potential target genes were predicted for 39 families and 11 targets were predicted for 4 among the 17 identified novel microRNAs. For understanding better the functions of microRNAs, target genes were analyzed by Gene Ontology and their pathways illustrated by KEGG pathway analyses. The presence of microRNAs identified in jute was validated by stem-loop RT-PCR followed by end point PCR and qPCR for randomly selected 20 known and novel microRNAs. This study exhaustively identifies microRNAs and their target genes in jute which will ultimately pave the way for understanding their role in this crop and other crops.

1. Introduction

Regulation of gene expression is one of the most enigmatic facets of molecular genetics that results in intricate appearance of a biological entity. Scientists have been attempting to elucidate the regulatory mechanisms of gene expression for long and the radical discovery of regulatory function of endogenous small noncoding RNAs is overwhelming the scientific community with their ever increasing potentials [1]. Small noncoding RNAs of 18–40 nucleotides (nt) in size have been proved to play a vital role in a remarkably wide range of biological processes, including cell proliferation, developmental timing and patterning, chromatin modification, genome rearrangement, and stress response in plants and animals [2]. Small RNAs regulate a variety of biological processes in plants by interfering with messenger RNA (mRNA) translation, directing mRNA cleavage or promoting the formation of compact, transcriptionally inactive chromatin

[3]. Several distinct classes of small RNAs have been reported so far including microRNAs (miRNAs) [4], small interfering RNAs (siRNAs) [5], repeat-associated small interfering RNAs (ra-siRNAs) [6], Piwi interacting RNAs (piRNAs) [7], natural antisense transcript derived small interfering RNAs (nat-siRNAs) [8], transacting small interfering RNAs (ta-siRNAs) [9], heterochromatic small interfering RNAs (hc-siRNAs) [10], secondary transitive small interfering RNAs, primary small interfering RNAs, competing endogenous RNAs (ceRNAs), and long small interfering RNAs [11–13]. It has only been a few years since it was appreciated that microRNAs provide an unanticipated level of gene regulation in both plants and metazoans [11].

miRNAs are well differentiated due to some of their particular characteristics [12]; they are derived from distinct genomic loci and processed from transcripts that can form local RNA hairpin structures, and usually miRNA sequences are nearly always conserved in related organisms [13, 14].

Most miRNAs are transcribed by RNA polymerase II which folds into a stable, usually imperfect, hairpin structure [13]; pri-miRNA transcript is cleaved to pre-miRNA by RNaseIII-type Dicer-like 1 (DCL1) protein [15] to produce a distinctive ~21 nt, double-stranded RNA. This duplex is exported into the cytoplasm by HASTY and methylated at the 3' end by HEN1 [16]. A cytoplasmic helicase unwinds the translocated duplex into a single-stranded mature miRNA, which is finally incorporated into RNA-induced silencing complex (RISC) [5, 17, 18]. A mature miRNA sequence can range from 19 to 24 nucleotides (nt) in length and act as a regulatory molecule in posttranscriptional gene silencing by base pairing with target mRNAs [12]. Within the RISC complex, miRNAs function in the direct cleavage of 3' untranslated region of protein-coding genes or translational repression depending on its perfect or imperfect match with the targets [19]. The same mature miRNA can also be present as several length variants; these populations of miRNA variants are called isomiRNAs, which are isoforms of microRNAs caused by an imprecise or alternative cleavage of Dicer during pre-miRNA processing [20].

Several miRNAs have been identified in plants, and they have been characterized in a wide variety of metabolic and biological processes with important functions [12]. The first plant miRNAs were described in *Arabidopsis thaliana* [21]; currently the latest miRBase release (v20, June 2013) contains 24,521 microRNA loci from 206 species, processed to produce 30,424 mature microRNA products [22]. Earlier, miRNAs have been identified through either bioinformatics analysis or sequencing [23]; various methods have been used to identify miRNAs in rice [24], wheat [25], tomato [26], and maize [27]. Besides the miRNAs that are highly conserved in different species, there are species-specific miRNAs originating from recently evolved miRNA genes [28, 29]. The expression of these species-specific miRNAs is often low and can therefore be difficult to detect by traditional methods [30]. In recent times, high-throughput sequencing platforms are showing significant promise for small RNA discovery and genome-wide transcriptome analysis at single-base pair resolution [23, 31]. Sequencing techniques such as the Solexa Platform, SOLiD, and 454 Technology as well as other massively parallel sequencing strategies have been successfully applied in order to identify miRNAs in many plant species, such as rice [32], alfalfa [33], grape [34], tomato [35], orange [36], soybean [37], peanut [38], poplar [39], and black gram [40]. In comparison with microarray, deep sequencing has several advantages, the major one being its application in comprehensive identification and profiling of small previously unknown RNA populations [23]. Nevertheless, analyses of these data are not perfect, especially in the absence of native genome sequence [40].

Jute (*Corchorus olitorius* and *Corchorus capsularis*) is a bast fibre, like flax and hemp. Cultivation of this environmentally friendly as well as the most affordable fibre producing plant is concentrated around the Ganges Delta region of Bangladesh and India where the warm, wet climate during the monsoon season provides ideal growing conditions. In terms of usage, production, and global requirement, jute is second only to cotton [41]. Jute plants are easy to grow, have

a high yield per acre, and, unlike cotton, have little need for pesticides and fertilizers. They are also known to enrich the soil [42]. As these plants grow fast, they are often used in crop rotation. The leaves and roots left after harvest enrich the soil with micronutrients, maintaining soil fertility. When used as a geotextile, it puts nutrients back in the soil when it decomposes. This rain-fed crop during its growth helps to clean the air by assimilating three times more CO₂ than an average tree, converting the CO₂ into oxygen.

Despite its great agronomic importance, research on jute at the molecular level is insignificant [43]. Genome sequence of jute is not available in the public database. So far only 1,210 sequences are found in the GenBank [43] with no deposits of any miRNA sequences in miRBase database. Within this context, the current study has employed the deep sequencing strategy in an attempt to effectively identify conserved and novel jute miRNAs. Quantitative real-time PCR (qRT-PCR) has been performed to determine the expression of these miRNAs. For mapping the identified miRNAs, genome of *Vitis vinifera* was used as reference because of sequence homology of this species with jute [44].

2. Methods and Materials

2.1. Plant Materials Preparation and Small RNA Library Construction. Seeds of farmer popular O-9897 variety of tossa jute (*Corchorus olitorius*) were collected from Bangladesh Jute Research Institute (BJRI). They were surface sterilized with 70% ethanol, subsequently washed in distilled water, and allowed to germinate on sterile petri dishes containing 3 MM moist filter paper (Whatman) at 30 ± 1°C and 65% relative humidity. Seeds were allowed to grow for 4 days under the specified conditions. On the fourth day of germination, seedlings were collected and immediately snap-frozen in liquid nitrogen and stored at -80°C for subsequent use.

RNA was isolated from collected seedlings using TRIzol reagent (Invitrogen, USA) by following the manufacturer's instructions. Later, RNA samples were sent to Beijing Genome Institute (BGI, Shenzhen, China) for deep sequencing of small RNA by Illumina Hiseq high-throughput sequencing platform. In short, the sRNAs pool of 18–30 nt in length were fractionated from total RNA. After ligation with 5' and 3' adaptors, the short RNAs so obtained were reverse-transcribed to cDNA according to the Illumina protocol. The resulting small RNA library was then sequenced following SBS method (sequencing by synthesis) by Illumina Hiseq high-throughput sequencing.

2.2. Prediction of Known miRNA in Jute. Raw data obtained from Illumina Hiseq high-throughput sequencing was at first filtered by removing contaminants which include low quality reads, reads with 5' primer contaminants, reads without 3' primer, reads without the insert tag, reads with poly A, and reads shorter than 18 nt. After cleaning, the final reads were then used for further analyses. Clean reads fully matching other RNAs, including mRNA, rRNA, tRNA, snRNA,

snoRNA, and repeat RNA, were excluded by using BLASTn-short alignment (blast2.2.26+, <ftp://ftp.ncbi.nih.gov/blast/executables/blast+/2.2.26/>) and aligning against Sanger RNA family database (Rfam 11.0, <ftp://ftp.sanger.ac.uk/pub/databases/Rfam>). The remaining unique sequences were further aligned against miRBase-v20 [22] allowing up to 3 mismatches to identify known miRNAs present in *C. olearius*. Mature miRNAs present within a genome encoding identical or nearly identical sequences were then grouped together into a family.

2.3. Novel miRNA Identification. Prediction of novel miRNA was done using prediction software, Mireap (<http://sourceforge.net/projects/mireap/>), developed by BGI by taking into consideration secondary structure, cleavage position of Dicer protein, and minimum free energy of the unannotated small RNA tags. Strategic conditions for selecting unique miRNA are as follows: (i) the tags which are to be used to predict novel miRNA should be from unannotated tags which can match reference genome (*Vitis vinifera*), from the tags which align to introns as well as antisense exons; (ii) genes, whose sequences satisfy the above standards and their secondary structures which allow hairpin miRNAs to fold within them and the presence of mature miRNAs in one arm of the hairpin precursors, are considered as candidate genes for miRNA; (iii) the possible candidate mature miRNA strand contains 2-nucleotide 3' overhang; (iv) hairpin precursors of the candidate miRNAs are devoid of large internal loops or bulges; (v) secondary structures of the hairpins are stable, with the minimum folding energy (MFE) lower than or equal to -20 kcal/mol.

2.4. Target Gene Prediction. The rules used for target prediction in plants are based on those suggested by Allen et al. [9] and by Schwab et al. [45]. These are (i) presence of maximum 4 mismatches between small RNA and target (G-U bases count as 0.5 mismatches), (ii) not more than 2 contiguous mismatches in the miRNA/target duplex, (iii) no end-to-end mismatches at the 5' of miRNA from 2 to 12 positions of the miRNA/target duplex, (iv) no mismatches in positions 10-11 of miRNA/target duplex, (v) a maximum of 2.5 mismatches in positions 1-12 of the miRNA/target duplex from 5' region of miRNA, and (vi) minimum free energy (MFE) of the miRNA/target duplex which should be $\geq 74\%$ of the MFE of the miRNA bound to its perfect complement. The targets of miRNAs were further validated by a well-recognized miRNA-target prediction tool: psRNA Target [46]. In addition to potential target prediction for known and novel miRNAs, pathways which include the corresponding target genes as well as the biological function of such genes are taken into consideration by using the grape genome as a reference. Biological functions are recommended by using GO (level 3) (<http://www.geneontology.org/>). Gene Ontology (GO) is an international standard classification system for gene function, which provides a set of controlled vocabulary to comprehensively describe the property of genes and gene products [47]. There are 3 ontologies in GO: biological process, cellular component, and molecular

function, containing lists of biological functions that illustrate each gene and its product (<http://www.geneontology.org/>). Each category defines precise participation of a given gene within an organism. As for pathway identification, KEGG (<http://www.genome.jp/kegg/>) [48] database was used to culminate the target genes within the systematic biological pathways. KEGG analyses reveal the main pathways with which the target gene candidates are involved [49].

2.5. Validation of the Presence of Jute miRNAs. To verify the identified known and potential novel miRNA candidates in jute, stem-loop reverse transcription-PCR was performed [50].

Stem-loop primers were designed according to the method described by Chen et al. [51]. This primer binds to specific miRNA at the 3' region owing to the precision conferred by the primer with the exact reverse complement of six nucleotides from the 3' end of each particular miRNA sequence, which is reverse-transcribed by the RT enzyme. Two thousand nanograms of total RNA were used to perform the RT reaction with Superscript III First Strand Synthesis System (Invitrogen, USA) according to the protocol of Varkonyi-Gasic et al. [50], which has been further standardized for jute. For a reaction volume of $20 \mu\text{L}$, $0.5 \mu\text{L}$ of 10 mM dNTPs was at first taken together with an appropriate amount of RNA and an adjusted amount of DEPC-treated H_2O followed by heating the mixture for five minutes at 65°C and then immediately transferring the same on ice. While keeping on ice for approximately 2 minutes, $4 \mu\text{L}$ of 5X FS buffer, $2 \mu\text{L}$ of 0.1 M DTT, $1.2 \mu\text{L}$ of $1 \mu\text{M}$ stem-loop primer, $0.25 \mu\text{L}$ of M-MLV Superscript III RT ($200 \text{ U}/\mu\text{L}$), and $0.1 \mu\text{L}$ RNaseOUT ($40 \text{ U}/\mu\text{L}$) were added to the reaction mix. The RT reaction was carried out in a thermal cycler (Mastercycler, Eppendorf, Germany), followed by a pulse RT cycle starting from incubation at 16°C for 30 minutes, then 60 cycles of 30 sec. at 30° , 30 sec. at 42°C , and 1 sec at 50°C . This step was followed by another incubation step of 5 mins at 85°C to inactivate the RT enzyme.

End point PCR was then conducted with miRNA specific forward primer and a universal reverse primer to check the presence of the specific miRNAs. PCR products were electrophoresed on 3% agarose gel in 1X TAE and stained with ethidium bromide before visualization under a transilluminator. We further conducted quantitative real-time PCR for confirming the expression of some selected known and novel miRNAs using a 32-well plate Roche LightCycler Nano System and the Roche SYBR Green Master I (Roche Diagnostics, Germany). Briefly, equal amount of cDNA was taken in a reaction volume of $7.5 \mu\text{L}$ in triplicate with $0.1875 \mu\text{L}$ of each primer (forward and universal reverse) and $3.75 \mu\text{L}$ of SYBR Green Master I. Thermo cycling conditions were set at an initial polymerase activation step for 600 seconds at 95°C , followed by 45 cycles of 5 sec at 95°C for template denaturation, 10 sec at 60°C for annealing, and 1 sec at 72°C for extension and fluorescence measurement. Later, a dissociation protocol with a gradient from 50°C to 95°C was used for each primer pair to verify the specificity of the RT-qPCR reaction and the absence of primer dimers.

TABLE 1: Summary of data cleaning.

Type	Count	Percent (%)
Total_reads	16912862	
High_quality	16822412	100%
3' adapter_null	12978	0.08%
Insert_null	1934	0.01%
5' adapter_contaminants	86857	0.52%
Smaller_than_18_nt	74711	0.44%
Poly A	1608	0.01%
Clean_reads	16644324	98.94%

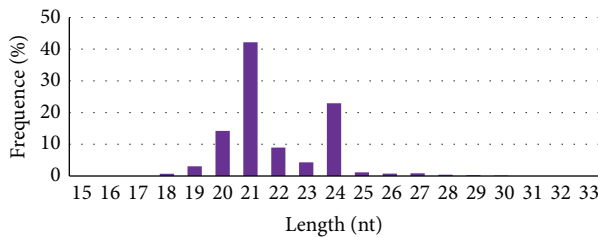


FIGURE 1: Length distribution of small RNAs found in jute (violet bar indicates the percentage of total tags).

3. Results

3.1. Deep Sequencing of Jute Small RNAs. In order to identify microRNAs in jute, RNA was isolated from the total tissue of four-day seedlings and subjected to Illumina Hiseq high-throughput sequencing by synthesis (SBS) technology. Among a total of 16912862 raw reads, 16822412 high quality reads were filtered through a series of data cleaning processes. The details of tag cleaning are summarized in Table 1, which shows that a total of 16644324 clean reads were obtained by removing 3' adapter null, insert null, 5' adapter contaminants, sequences smaller than 18 nt, and poly A.

This comprises about 99% of the high quality reads. Distribution of these clean reads that contain a pool of small RNAs ranging from 18 to 30 nucleotides is shown in Figure 1. However, the sizes of small RNAs were not found to be uniform; majority (92.69%) of the sRNAs are 20–24 nt in size, with 21 nt being the most abundant (42.19%), followed by 24 nt (22.95%) and 20 nt (14.25%), respectively. These sequences were then aligned to Rfam 11.0 [52] and Genbank database (BLASTn) to identify common sRNAs other than miRNAs, as well as to remove mRNAs (see supplementary file-1 in Supplementary Material available online at <http://dx.doi.org/10.1155/2015/125048>). The remaining sequences were matched against miRBase-20 database [53] for the prediction of miRNAs, revealing 33433 unique and 8994892 redundant reads, which were finally used to identify known miRNAs. The novel miRNAs in jute were identified from unannotated tags by using Mireap, software developed by BGI (described in Section 2).

3.2. Identification of Known miRNAs in Jute. In the absence of the complete genome sequence and with practically no

information on miRNA of jute in miRBase, clean reads were aligned to the miRNA precursor/mature miRNA of all plants in miRBase allowing up to three mismatches or free gaps [54] to identify known miRNAs. The expression of miRNA is generated by summing the count of tags which can align to the temporary miRNA database, generated by choosing the most expressive miRNA of each mature miRNA family.

A total of 227 known miRNAs were identified in this study, of which 164 belong to 23 conserved and 63 to 58 nonconserved families. These nonconserved families were further categorized into 18 defined and 40 undefined families. Conservancy of miRNA families found in jute showed high homology with their respective homologs in other model plants (Figure 2). However, the number of family members of conserved miRNAs was highly variable, with miR156 being the largest family, consisting of 26 members, whereas miR403, miR394, miR827, miR477, and miR2111 were the smallest among the families, comprising only one member. miR166 and miR169 were the second largest with each having 18 members and miR171 was the third largest family with 14 members (Figure 3). Most of the conserved families contain both 5p and 3p mature miRNA sequences, attaching a high confidence to the data set [54]. Highly variable reads numbers were also found among the families, even in members of the same family, indicating different expression levels of these miRNAs. Among them, col-miR157a had the highest level of expression, having 5531609 counts, and the other miRNAs like col-miR156a, col-miR166a, and col-miR167h also had relatively high reads numbers, counting more than 150000. Several conserved miRNAs (like miR171, miR398, and miR159) and, as expected, most of the nonconserved miRNAs had relatively low copy numbers. Interestingly, a miRNA named col-miR3954 from an undefined family had very high level of expression, having 868222 reads, third highest of all miRNAs found in jute. miRNAs from each family with highest reads number are shown in Table 2. High expression frequency of miRNAs derived from the 3' arm of some pre-miRNAs, like col-miR166h-3p, col-miR166g-3p, col-miR166j-3p, col-miR165a-3p, col-miR396b-3p, col-miR396e-3p, and so forth, compared to their corresponding 5' arm-miRNAs, supports the observations of functional activity of both arms of pre-miRNA hairpins [55, 56]. Details of the miRNAs found in jute are summarized in supplementary file-2.

3.3. Identification of Novel miRNAs. The miRNA hairpins are mostly located in intergenic regions, introns, or reverse repeat sequence of coding sequences [57]. Thus tags belonging to these regions were used to predict novel miRNAs. Characteristic hairpin structure of miRNA precursor was used to predict novel miRNA with prediction software Mireap (<http://sourceforge.net/projects/mireap/>) by exploring the secondary structure, the Dicer cleavage site, and the minimum free energy of the unannotated small RNA tags which could be mapped to the reference *Vitis vinifera* genome. Predicted secondary structures were further validated by Mfold (supplementary file-3) [58] and novel miRNAs were identified based on the selection criteria described in Section 2. 17 potential novel miRNAs have been identified in this study of

TABLE 2: miRNAs from each family with the highest frequency in jute with their homologs in other plants.

Jute family	Jute miRNA	Length of small RNA sequence	Count of small RNA sequences	Small RNA sequence	Homolog (best one)	Matches number	Mismatches number	Gaps number
Conserved								
miR156	col-miR157a	21	5531609	UUGACAGAAGAUAGAGAGCAC	ath-miR157a	21	0	0
miR166	col-miR156a	20	1724997	UGACAGAAGAGAGUAGAGCAC	ath-miR156a	20	0	0
miR167	col-miR166a	21	215636	UGGACCAGGGCUCAUUCGCC	ath-miR166a	21	0	0
miR396	col-miR167h	22	154973	UGAAGCUGCCAGCAUUCUUA	mdm-miR167h	22	0	0
miR399	col-miR396b-3p	21	18695	GCUCAAAGAAAGCUGUGGGAGA	gma-miR396b-3p	21	0	0
miR168	col-miR168a	21	16590	UCGCUUGGUGCAGGUCGGGAA	ath-miR168a	21	0	0
miR164	col-miR164a	21	7491	UGGAGAAGCAGGGCACGUGCA	ath-miR164a	21	0	0
miR169	col-miR169b	21	2528	CAGCCAAGGAUACUUGCCGG	ath-miR169b	21	0	0
miR390	col-miR390a	21	2952	AAGCUCAGGAGGGAUAGCGCC	ath-miR390a	21	0	0
miR160	col-miR160a-3p	21	1448	GCGUAUGAGGAGCCAAAGCAUA	gma-miR160a-3p	21	0	0
miR159	col-miR159a	21	484	UUUGGAUUGAAGGGAGCUCUA	ath-miR159a	21	0	0
miR171	col-miR171b	21	558	UGAUUGAGCCUGGCCAAUAUC	osa-miR171b	21	0	0
miR403	col-miR403	21	282	UUAGAUUCACGCACAACUCG	ath-miR403	21	0	0
miR398	col-miR398	21	262	GGAGCGACAUGAGAUACACAU	hbr-miR398	20	1	0
miR482	col-miR482b	22	2206	UCUUACCUACUCCACCCAUGCC	ghr-miR482b	21	1	0
miR408	col-miR408	21	129	AUGACUGCCUUCUCCUUGGC	ath-miR408	21	0	0
miR397	col-miR397a	21	106	UCAUUGAGUGCAGCGUUGAUG	ath-miR397a	21	0	0
miR530	col-miR530a	21	81	UGCAUUUGCACCCUGCACCUUU	csi-miR530a	20	1	0
miR393	col-miR393b-3p	21	61	AUCAUGCGAUCCUUCGGAAU	stu-miR393-3p	20	1	0
miR394	col-miR394a	20	16	UUGCAUUCUGUCCACCUCC	ath-miR394a	20	0	0
miR827	col-miR827a	21	22	UUAGAUGACCAACAAACA	ghr-miR827a	21	0	0
miR477	col-miR477i	21	2	ACUCCUCCUCAAAGGCUUCCG	mes-miR477i	21	0	0
miR211	col-miR211a	21	15	UAAUCUGCAUCCUGAGGUUUUG	ptc-miR211a	21	0	0
miR172	col-miR172a	21	1480	AGAAUCUUGAUGAUGCUGCAU	ath-miR172a	21	0	0
Nonconserved								
miR2275	col-miR2275a-3p	22	54	UUAAGUUUUUCCCAUAUCUCA	zma-miR2275a-3p	20	1	2
miR2118	col-miR2118a-3p	22	36	UUGCCGAUCCGCCCAUUCGGU	gma-miR2118a-3p	19	2	1
miR528	col-miR528-5p	21	17	UGGAAAGGGCAUUCGAGAGGAG	osa-miR528-5p	21	0	0
miR1310	col-miR1310	21	69	AGGCAUCGGGGCGCAACGCC	han-miR1310	21	0	1
miR7696	col-miR7696a-3p	21	6	UCUGAAUCAUGAGAACUUGAG	mtr-miR7696a-3p	19	1	2
miR6224	col-miR6224a-3p	21	4	CUGAAUAUAUAGGACGGGGG	sbi-miR6224a-3p	19	1	2
miR6462	col-miR6462c-5p	21	28	AAGGGACAAAAGGCUUAUAG	ptc-miR6462c-5p	20	0	3
miR4243	col-miR4243	21	3	UUGAACUUUGAACGAUUUCGAC	ath-miR4243	19	1	2
miR5745	col-miR5745b	21	16	UUUAAUUUAUAUCAUCUCAC	mtr-miR5745b	20	0	2
miR902	col-miR902j-5p	24	1	AUAUGUACGCAGAUUCUUAUUU	ppt-miR902j-5p	21	0	3
miR2873	col-miR2873b	21	10	UUUGGCGUGAGUUUGGUUUG	osa-miR2873b	19	1	2
miR2950	col-miR2950	21	2465	UGGUGUGCAGGGGUGGAAUA	ghr-miR2950	21	0	0
miR5067	col-miR5049c	24	3837	GGACAAUUUUGGGACGGAGGG	hvu-miR5049c	21	2	1
miR818	col-miR1436	23	920	AGAUAAUUGGGACGGGGAGU	osa-miR1436	20	1	2
miR4414	col-miR4414a-3p	21	11	AUCCAACGAUUCGAGGAGCUCG	mtr-miR4414a-3p	20	1	0

TABLE 2: Continued.

Jute family	Jute miRNA	Length of small RNA sequence	Count of small RNA sequences	Small RNA sequence	Homolog (best one)	Matches number	Mismatches number	Gaps number
miR1509	col-miR7122a	22	488	UUGGACAGAGAAAUCACGGGUGG	mdm-miR7122a	20	2	0
miR158	col-miR158a	20	34	UCCCAAAUUGAGACAAAGCA	ath-miR158a	20	0	0
miR161	col-miR161.2	21	2	UCAUUGCAUUGAAAGUGACUA	ath-miR161.2	21	0	0
Undefined	col-miR5162	24	4	AAAUGACCAAAUACCCCUAAAU	osa-miR5162	22	1	2
Undefined	col-miR6248	20	1	UAAUUGAGGAUGGAGGGAGU	osa-miR6248	18	2	1
Undefined	col-miR7767-3p	22	1	UAGGAUCAGGCAGCUUGAAGGU	bdi-miR7767-3p	19	2	1
Undefined	col-miR5997	21	1	UGAAACUCAAGUAGCUAAAAG	ath-miR5997	20	0	2
Undefined	col-miR5057	23	1	AAACUUUCAGAUUGCAUUUGACA	bdi-miR5057	20	1	2
Undefined	col-miR6172	21	1	UGAGACCCUUGUUAAGUUAGAA	hbr-miR6172	19	1	2
Undefined	col-miR6279	20	3	UAACAAGAAUUCAGACACA	ppe-miR6279	18	2	1
Undefined	col-miR6220-3p	23	1	AGACUUAAUUUGGACGGGAGA	sbi-miR6220-3p	21	2	1
Undefined	col-miR6443	21	1	UGUAUGAUCAUAGUGCUGGAG	ptc-miR6443	19	1	2
Undefined	col-miR156h	20	17	UGACAGAAGAGAGAGAGCAU	vvi-miR156h	20	0	0
Undefined	col-miR3954	22	868222	UUGGACAGAGUUAUCACGGGUGG	csi-miR3954	19	2	1
Undefined	col-miR1671	20	1	UCAUGCUGGCAGCUUCACUU	gma-miR1671	20	0	3
Undefined	col-miR6300	19	4400	GUCGUUUGUAGUAGUGGU	gma-miR6300	18	0	1
Undefined	col-miR169p	21	6	UAGCCAAGGACAAUUGCCGG	osa-miR169p	21	0	1
Undefined	col-miR894	20	5406	GUUUCACGUCGGGUUCACCA	ppt-miR894	19	0	2
Undefined	col-miR472a	22	218	UUUUCUACUCCUCCAUCCG	ptc-miR472a	21	1	0
Undefined	col-miR5059	21	212	CGGUCUUGGACAGCAACCA	bdi-miR5059	19	1	1
Undefined	col-miR2916	22	312	GGGGUCUGAAAGCAUCAGAU	peu-miR2916	20	2	1
Undefined	col-miR5072	21	90	CGUCCCAAGGAGUCGCCA	osa-miR5072	21	0	1
Undefined	col-miR477h	22	4	ACUCUCCUCAAAGGCUUCCAG	mes-miR477h	21	0	1
Undefined	col-miR6478	21	164	CCGACCUUAGCUCAGUUGGUA	ptc-miR6478	20	1	0
Undefined	col-miR5205b	24	306	CUUAUAAUAGGACAGAGGGAGU	mtr-miR5205b	23	1	0
Undefined	col-miR7505	21	31	UUCAGAAAACCAUCCCUCCUU	ghr-miR7505	20	1	0
Undefined	col-miR5077	20	378	GAUUCACGUCGGGUUCACCA	osa-miR5077	18	1	1
Undefined	col-miR5054	20	904	GUUCCCCACAGUCGGGCCA	bdi-miR5054	17	1	2
Undefined	col-miR845c	24	28	AGGUCUCUAUACCAUUGACGUAG	vvi-miR845c	21	0	3
Undefined	col-miR916	22	28	CGAAGGUCUGCGGUCCGAAUCC	cre-miR916	19	2	1
Undefined	col-miR161-5p.1	21	11	UUGAAAUGACUAUCUGGGG	aly-miR161-5p.1	21	0	0
Undefined	col-miR1863	24	56	AGCUCUGAUACCAUUGUUAAGCAUC	pab-miR1863	21	1	2
Undefined	col-miR7490	24	41	AGUCUGAUAAACUCCACUGACGGU	ghr-miR7490	22	1	2
Undefined	col-miR3946	21	13	UUGAGAGAAAGAGAGAGCAC	csi-miR3946	21	0	3
Undefined	col-miR7728-5p	19	5	UUCGGAUUGAGUGGAUUUU	bdi-miR7728-5p	18	1	2
Undefined	col-miR1862f	21	1	AAGGGUUGGUUUAUUUUUGG	osa-miR1862f	18	2	1
Undefined	col-miR6171	21	11	ACUAGGAUUGCUGAAAGUUU	hbr-miR6171	19	1	2
Undefined	col-miR5021	21	31	UAAGAAAGAAUAGAAGAAUAA	ath-miR5021	18	2	1
Undefined	col-miR5244	21	6	UAUCUGAUGAUGAUUUUGGU	mtr-miR5244	19	2	0
Undefined	col-miR5049-3p	23	30	AAGUAAUUGGAACGGAGGGAGU	bdi-miR5049-3p	21	2	1
Undefined	col-miR5629	23	41	UUAGGGUAGUUAACGGGUAGUUA	ath-miR5629	21	1	1

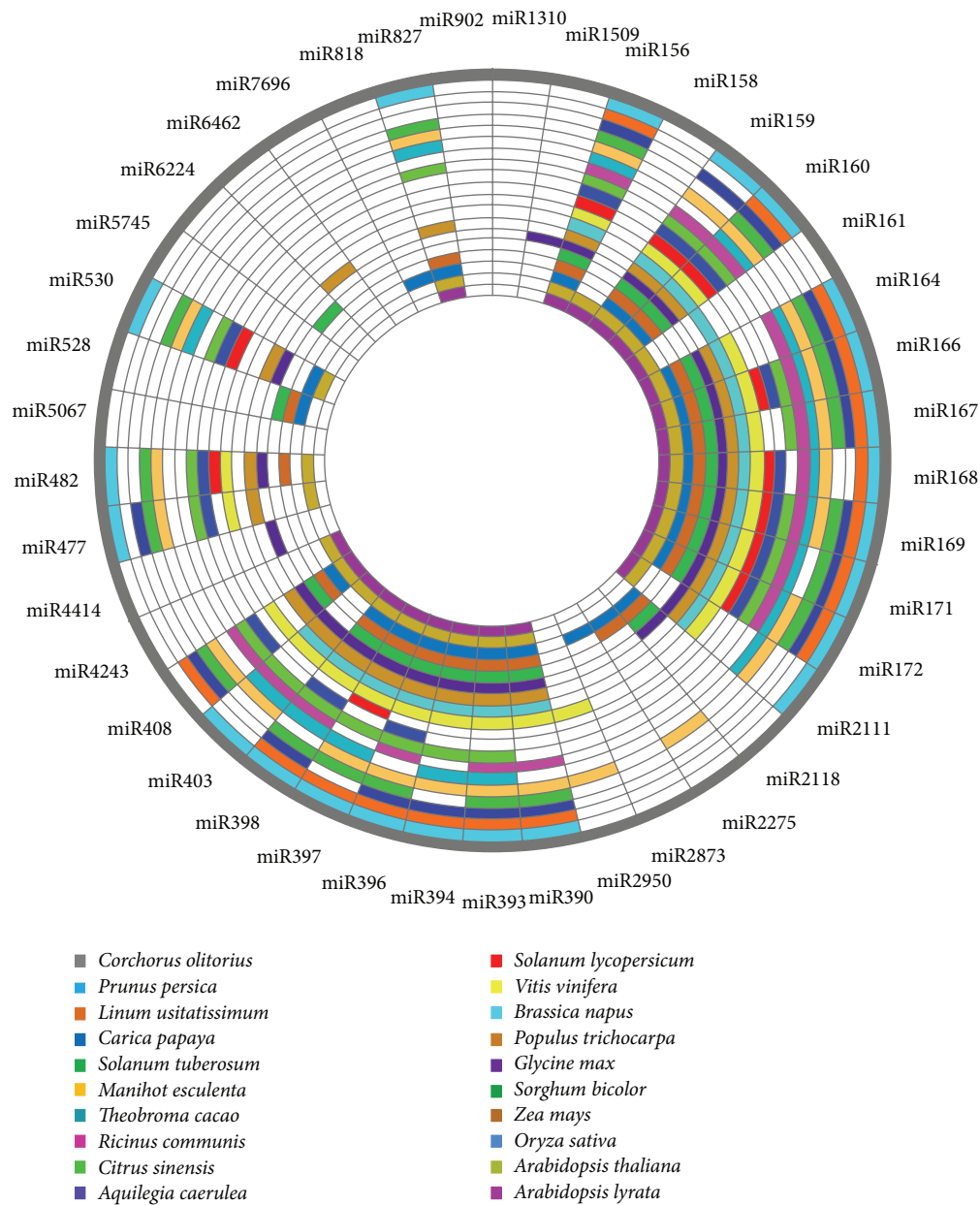


FIGURE 2: Conservancy of miRNAs identified in jute is presented as circular heat map among different model plants. Each color represents a different plant species and white color represents absence of miRNA. miRNA that was found in at least 9 plants was considered as conserved.

which 9 miRNAs were derived from 3' arm of the pre-miRNA sequences and 8 from the 5' arm (Table 3). Average length of the pre-miRNAs sequences ranged from 78 to 349 nt, similar to those found in maize [59] and rice [24]; minimum folding energy (MFE) for jute miRNAs was observed to be within a range from -21 to -105.3 kcal/mol, similar to the range observed in cucumber [60] (supplementary file-4). Expression of novel miRNA was determined by summing the count of such miRNAs which have no more than 3 mismatches on either the 5' or 3' ends and with no mismatch in the middle. Novel miRNAs usually have lower levels of expression than the conserved miRNAs, as evident from findings of several plant species like soybean, *Brassica napus*, maize, *Arabidopsis*, and wheat [59, 61–64].

3.4. Target Gene Prediction for Identified miRNAs. For a precise elucidation of the role of miRNAs, target identification and determination of their biological functions are of vital importance. With a plethora of experimentation, it is now evident that cleavage or translational repression site of most known plant miRNAs is located in the CDS (coding sequence) region of their target mRNA with perfect or nearly perfect sequence complementarity [65], making it feasible to identify plant miRNA targets [4, 21, 66]. In this study, target genes of miRNAs were identified by BLASTn against the genome sequence of *Vitis vinifera*, following methods described by Allen et al. and Schwab et al. [9, 45]. Among a total of 79 identified miRNA (both conserved and nonconserved) families, 116 potential target genes were

TABLE 3: Potential novel miRNAs found in jute.

miRNA name	Number of reads	Mature miRNA sequence	Mature miRNA length	MFE (kcal/mol)
col-miRN1-5p	1448	GUGGGCGUGCCGGAGUGGUUA	21	-28.9
col-miRN2-3p	219	AGAGGGACUAUGGCCGCUUA	20	-53.5
col-miRN3-3p	17	UCGGUUUUGAAUUAGAGACGU	21	-85
col-miRN4-3p	14	UGAUGAUUGUGAAGAAGAUGA	21	-66.34
col-miRN5-3p	32	AGAGGCUCGGUGAAAUAGACAU	22	-24.62
col-miRN6-5p	11	UUCGUCCCCGGCAACGGCGCCA	22	-66.6
col-miRN7-5p	7	UUUUUUAAUUUUUUUAUUUAUC	21	-21
col-miRN8-5p	20	GUUGAUCAAGUUGUGGAUGGC	21	-79.32
col-miRN9-3p	2	AAACUUCGAAUUGGGAGGGC	20	-89.3
col-miRN10-3p	3	UGAAUGAUUUCGGACCAGGCU	21	-48.3
col-miRN11-3p	2	GUAAGAAGGGGUAGAGAAAAU	21	-34.9
col-miRN12-3p	5	AAGAUAGAGAGCACAGAUGAU	21	-51.1
col-miRN13-5p	3	GGCGCUGCCUACUCACUCGGACA	23	-40.77
col-miRN14-3p	7	GUGAGGCUGGUUUCACAGAGCA	22	-39.1
col-miRN15-5p	6	GAGUGCAGCCAAGGAUGACUU	21	-64.9
col-miRN16-5p	4	UCAAGGUGGAGAUUGUUAGGA	21	93.4
col-miRN17-5p	6	UUUAUCGAUGUGGGUAUUAC	21	-105.3

TABLE 4: Target genes for jute specific miRNAs.

miRNA name	Targets number	Target accession	Annotation	Location	Free energy
col-miRN1	1	GSVIVT01015521001	Pentatricopeptide repeat-containing protein, mitochondrial	2244, 2264	-47.30 [100.00%]
col-miRN4	2	GSVIVT01020089001	Thioredoxin H	16, 36	-25.50 [77.74%]
		GSVIVT01021522001	Protease degS	67, 87	-24.50 [74.47%]
		GSVIVT01000651001	Conserved gene of unknown function	178, 198	-13.30 [76.00%]
		GSVIVT01000655001	NB-ARC domain containing protein	1291, 1311	-13.30 [76.00%]
col-miRN7	6	GSVIVT01000657001	NB-ARC domain containing protein	1246, 1266	-13.20 [75.86%]
		GSVIVT01021549001	Conserved gene of unknown function	702, 722	-12.90 [75.00%]
		GSVIVT01035288001	Casein kinase	1464, 1484	-14.20 [81.14%]
		GSVIVT01000656001	NB-ARC domain containing protein	1300, 1320	-13.30 [76.00%]
col-miRN8	2	GSVIVT01033994001	26S proteasome regulatory particle non-ATPase subunit 8	835, 855	-30.70 [78.52%]
		GSVIVT01037657001	Aconitase	30, 043, 024	-29.30 [77.11%]

predicted for 39 families (supplementary file-5). A total of 46 genes from this prediction overlapped with targets identified by psRNA Target which found 99 target genes for 19 miRNA families (supplementary file-9). Highest number (16) of targets was identified for miR397 family, all of which are laccase, an enzyme, involved in plant cell wall lignification [67]. miR3946 had the second highest number of targets with 11 genes. Most of the other families targeted only a single gene. For the novel jute miRNAs, a total of 11 targets were predicted for 4 among the 17 identified miRNAs (Table 4, details in supplementary file-6), with a maximum number of target genes (6) recognized for col-miRN7. Most of col-miRN7 targets are NB-ARC domain containing protein, which is

a resistance (*R*) protein, involved in pathogen recognition and subsequent activation of innate immune responses [68]. To better understand the functions of miRNAs, target genes were analyzed by Gene Ontology (GO) level 3 to divulge the regulatory network of miRNAs and target genes [69]. Such analysis demonstrates that for jute 133 predicted target genes (both for known and novel miRNAs) can be classified into 20 having biological, 5 cellular, and 5 molecular functions. Same gene was found to be involved in multiple processes with the reverse being also true (Figure 4 and supplementary file-7). As illustrated by KEGG pathway analysis (supplementary file-8) [70], the predicted target genes of jute miRNAs were found to be involved in 42 different pathways.

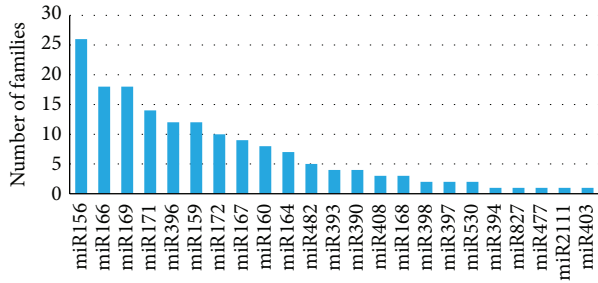


FIGURE 3: Number of family members of conserved miRNAs are represented as bar diagram.

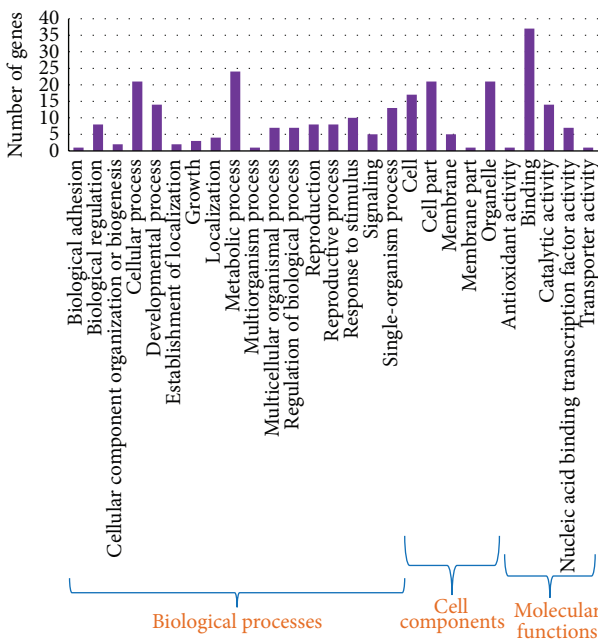


FIGURE 4: GO (level 3) annotation of predicted targets. Violet bar indicates the number of targets involved in each process.

3.5. Validation of the Presence of Known and Novel miRNAs in Jute. Some of the miRNAs identified through deep sequencing were verified by the standard stem-loop RT-PCR method [50] followed by end point PCR and qRT-PCR. The stem-loop primers were designed with a 3' specificity for a particular miRNA which hybridizes to the same and is reverse-transcribed by the RT enzyme. These primers increase the sensitivity of the reactions such that this method can significantly distinguish two miRNAs with only one single nucleotide change [51]. The RT product is then subjected to end point and qRT-PCR. Forward primers were precisely designed from the first 15 bases of each miRNA with 5' extension of random GC rich sequence to increase the melting temperature as mentioned by Varkonyi-Gasic et al. in 2007 while the reverse primer is a universal sequence that is designed from the 5' region of the stem-loop RT primer [71]. A set of 11 randomly selected conserved miRNAs as well as 9 novel miRNAs were used for verification. In this study, the stem-loop primer used was 50 bp long together with 5'

(forward primer) and 3' extensions, and the end point PCR product size ranged from 60 to 70 bp. Amplification of the product gave a sharp band for each of the selected known and novel miRNAs (shown in Figure 5). cDNAs were further amplified by qRT-PCR in technical triplicates from which log₂ values of C_q were calculated for each of the miRNAs and average of these values was compared with the log₂ value of read counts obtained from deep sequencing. Most of the qRT-PCR results acceded with the sequencing data; however, in some cases, discrepancy was observed (Figure 6).

4. Discussion

Widespread discovery of miRNAs and their critical role in gene regulation has made it ever important to recognize them in different species. Identification of miRNAs and their targets is the basis for understanding their physiological functions [60].

While a large amount of miRNAs are reported and deposited in databases from different plants, miRNA associated research in jute is still to be instigated. Without the genome sequence of jute at hand, identification of miRNA and their targets in jute by deep sequencing of small RNAs has been the greatest challenge of the current study. Use of closely related species' genomes as proxy references can facilitate miRNA identification in nonmodel species like jute for which genome sequence is not available [72]. We have used the grape genome as the background because of sequence similarity between these two species.

sRNAs with known function are commonly 20–24 nt in size [34]. Analyses of size distribution patterns of the reads show that the most abundant sRNAs in jute are 21 nt in size which is about 42.19%, consistent with recent identification of sRNAs in different plant [34, 62, 64].

Sequencing frequencies for miRNAs in a library can be used as an index for estimating the relative abundance of miRNAs [73]. Numerous small RNA sequences, engendered from Illumina Hiseq high-throughput sequencing platform, show the presence of different miRNA families and are even able to differentiate between distinct members of a given family. miR156 family which is highly conserved across the species [74] was found to be the largest family in jute seedlings with the highest expression of col-miR157a followed by col-miR156a. Two other members of the same family, namely, col-miR156c and col-miR156k, also show significant levels of expression. During shoot development, miR156 regulates the transition of plants from juvenile to adult phase by targeting SPL genes [75]. In *Arabidopsis*, miR156 is strongly expressed during seedling development and shows weak expression in mature tissues [76]. This could explain the relative abundance of the members of miR156 family since RNA used in sequencing was extracted from jute seedlings. Deep sequencing technology allows distinguishing and measuring miRNA sequences with only a few nucleotide changes [38]. Members of different families exhibit considerably dissimilar expression levels. For example, the abundance of miR156 family varied from 1 read (col-miR156p) to 5531609 reads (col-miR157a). This was also the case for some other miRNA

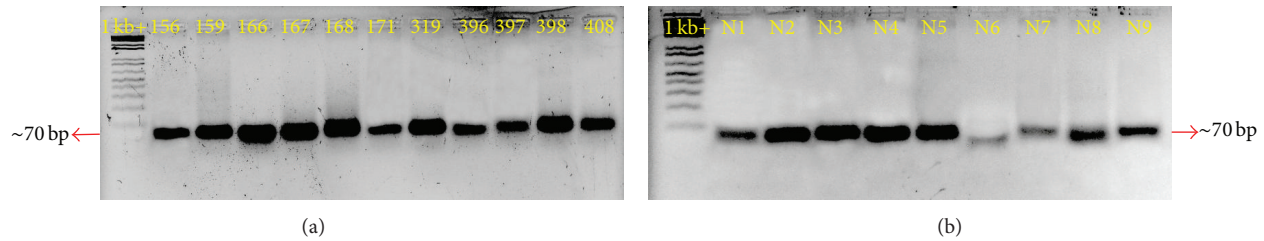


FIGURE 5: Agarose gel electrophoresis of known and novel miRNAs identified in jute. (a) Amplicons for known miRNAs. 156: miR156, 159: miR159, 166: miR166, 167: miR167, 168: miR168, 171: miR171, 319: miR319, 396: miR396, 397: miR397, 398: miR398, and 408: miR408. (b) Amplicons for novel miRNAs. N1: col-miRN1, N2: colmiRN2, N3: colmiRN3, N4: colmiRN4, N5: colmiRN5, N6: colmiRN6, N7: colmiRN7, N8: colmiRN8, and N9: colmiRN9.

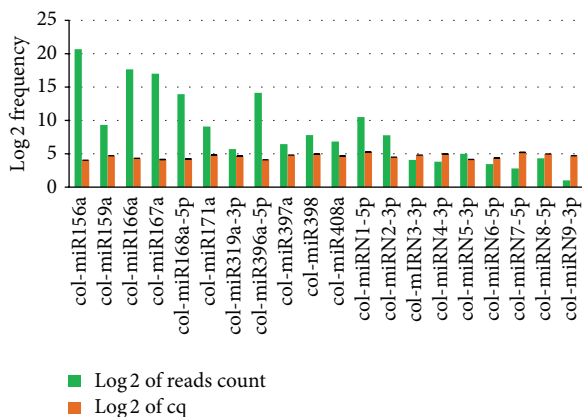


FIGURE 6: Comparative expression analysis of different selected miRNAs found by deep sequencing and qRT-PCR. Read counts of deep sequencing and cq values of qRT-PCR were converted into log₂ value for a better representation. Here the green bars represent log₂ values of sequencing frequency and orange bars represent the log₂ values of cq produced by qRT-PCR. Black regions on top of the orange bars represent errors calculated as standard deviation.

families, such as col-miR166 (from 3 to 215636 reads) and col-miR167 (from 12 to 154973 reads). Presence of a prevailing member in a miRNA family may indicate the dominant role of this member during the growth phase at which the samples were collected. It is also to be noted that most of the conserved miRNA families consist of more than one member, whereas nonconserved miRNAs identified in this study are mostly represented by a single MIR (miRNA) gene.

It has been hypothesized that MIR genes originate by gene duplication events followed by random mutation processes to evolve in multiples of imperfectly paired hairpins [77, 78]. Consequently, ancient evolutionarily conserved miRNAs are represented by multiple MIR genes whereas nonconserved miRNAs (believed to be evolutionarily recent) generally originate from a single locus [79]. It is plausible that the conserved miRNAs are responsible for control of basic cellular and developmental pathways common to most eukaryotes whereas nonconserved miRNAs are involved in regulation of species-specific pathways and functions [80].

Species-specific miRNAs are believed to have recently evolved and, in general, expressed at levels lower than those

of strictly conserved miRNAs [34, 77]. Data acquired from sequencing frequencies of conserved and nonconserved miRNAs fits well with this extrapolation, where the nonconserved and species-specific miRNAs show residual accumulation in the tested tissue. However, one miR-3954, a single member of an undefined family, appears to be expressed in significantly high levels. Its only homolog deposited in miRBase v20 is in *C. sinensis* [81], showing high frequency of reads. Though not deposited in miRBase, it has been reported in *X. sorbifolia* [82].

17 new jute specific miRNAs identified in this study show a size anticipated for sRNAs derived from DCL1 processing, although sequence variants that possess shortened or extended 5' or 3' ends were also found. Ten among the seventeen new col-miRNAs are 21 nt in size, consistent with canonical DCL1 products [79]. However length variation was also found. Two, col-miR2 and col-miR9, are 20 nt in size; three, col-miR5, col-miR6, and col-miR14, are 22 nt long. col-miR13 was found to be 23 nt in size which can probably be explained by the fact that diverse miRNA families are also independently processed by DCL3 to generate a new class of bona fide (23–25 nt) miRNAs with no canonical size, called long miRNAs [83].

A total of 20 miRNAs of both conserved and species-specific origin were corroborated by stem-loop RT-PCR and their expression pattern was assessed by qPCR to validate the data obtained from deep sequencing. Discrepancies in the expression pattern of some miRNAs found by deep sequencing and qPCR can be attributed to practical differences between the sensitivity and specificity of these two techniques [84]. The sensitivity and large dynamic range of next generation sequencing (NGS), along with its consistent prediction of fold changes when compared with gold-standard qPCR, support its use for discovery-oriented and exploratory miRNA profiling experiments [84, 85].

To evaluate and outline a putative function for a miRNA in plants, target identification is necessary [73]. We have predicted target genes for known and potential new miRNAs identified in this study using the genome of *Vitis vinifera* as a reference. Most of the target genes for conserved miRNA families predicted in jute have already been confirmed in model plants, as target genes are commonly conserved [78, 80]. miR156/157-Squamosa promoter-binding protein [86],

miR166-Homeodomain Leucine Zipper protein III (HD-ZIP III) [87], miR167-auxin response factor (ARF) [88], miR164-NAC domain protein [89], miR172-transcription factor APETALA2 [90], miR159-MYB transcription factor [91], miR171-GRAS family transcription factor [92], miR394-F-box family protein [93], and miR395-ATP sulfurylase [94] well characterized miRNA-target pairs in other plants have been found in jute. However, a number of widely studied miRNA-target pairs, such as miR398-copper superoxide dismutase [95], miR399-E2 ubiquitin conjugating protein [96], and miR162-Dicer-like 1(DCL1) [97], were not found in this study. This could possibly be due to the fact that the jute genome sequence is not available to be used as a reference. However, conserved miRNAs with their nonconserved targets, including miR167-peroxidase29, miR396-eukaryotic translation initiation factor 2c, miR168-NAC domain containing protein, miR164-growth regulating factor 1, miR390-AP domain containing transcription factor, miR160-MYB transcription factor, and miR393-GTP-binding protein alpha subunit, were also found to be present in jute, allowing presumption of nonconserved targets for conserved miRNAs. Highest number of target genes were identified for miR397, which is laccase, a well-studied enzyme, encoded by multigene families in poplar, *Arabidopsis*, rice, and *Liriodendron tulipifera* [98], reported to be involved in lignin biosynthesis of plants [99–101]. High lignin content of jute fibre limits its use in making fine fabrics [102]. Toughness of this biopolymer also poses a major obstacle to pulping, forage digestibility, and biofuel production [103]. It has been reported that transgenic *P. trichocarpa* plants overexpressing Ptr-miR397a result in a reduction of Klason lignin content [104], supporting the idea that use of miR397 would be an attractive means for reducing lignin-related problems. Future experiments including in-depth studies of miR397-laccase pair may help in producing quality products from jute.

5. Conclusion

This is the first report on jute miRNA identification. This set of experimentations for identification of miRNAs and their potential targets can initiate further study on understanding the mechanisms of regulation of jute miRNA.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Authors' Contribution

Md. Tariqul Islam and Ahlan Sabah Ferdous contributed equally.

Acknowledgments

The authors thank the Ministry of Science and Technology, Government of Bangladesh, for funding this project and Md. Moniruzzaman from LalTeer Livestock for shipment

of samples for sequencing. The authors acknowledge Arif Mohammad Tonmoy for his help and also appreciate the technical help from Md. Kamal Hossain.

References

- [1] J. Jiang, Y. Yang, and J. Cao, "Identification of microRNAs potentially involved in male sterility of *Brassica campestris* ssp. *chinensis* using microRNA array and quantitative RT-PCR assays," *Cellular and Molecular Biology Letters*, vol. 18, no. 3, pp. 416–432, 2013.
- [2] F. Xie, C. N. Stewart, F. A. Taki, Q. He, H. Liu, and B. Zhang, "High-throughput deep sequencing shows that microRNAs play important roles in switchgrass responses to drought and salinity stress," *Plant Biotechnology Journal*, vol. 12, no. 3, pp. 354–366, 2014.
- [3] B. Khraiweh, G. Pugalenti, and N. V. Fedoroff, "Identification and analysis of red sea mangrove (*Avicennia marina*) microRNAs by high-throughput sequencing and their association with stress responses," *PLoS ONE*, vol. 8, no. 4, Article ID e60774, 2013.
- [4] B. J. Reinhart, E. G. Weinstein, M. W. Rhoades, B. Bartel, and D. P. Bartel, "MicroRNAs in plants," *Genes & Development*, vol. 16, no. 13, pp. 1616–1626, 2002.
- [5] S. M. Hammond, E. Bernstein, D. Beach, and G. J. Hannon, "An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells," *Nature*, vol. 404, no. 6775, pp. 293–296, 2000.
- [6] A. Djikeng, H. Shi, C. Tschudi, and E. Ullu, "RNA interference in *Trypanosoma brucei*: cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs," *RNA*, vol. 7, no. 11, pp. 1522–1530, 2001.
- [7] A. A. Aravin, G. J. Hannon, and J. Brennecke, "The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race," *Science*, vol. 318, no. 5851, pp. 761–764, 2007.
- [8] V. N. Kim, "Sorting out small RNAs," *Cell*, vol. 133, no. 1, pp. 25–26, 2008.
- [9] E. Allen, Z. Xie, A. M. Gustafson, and J. C. Carrington, "microRNA-directed phasing during *trans*-acting siRNA biogenesis in plants," *Cell*, vol. 121, no. 2, pp. 207–221, 2005.
- [10] B. J. Reinhart and D. P. Bartel, "Small RNAs correspond to centromere heterochromatic repeats," *Science*, vol. 297, no. 5588, p. 1831, 2002.
- [11] D. V. Dugas and B. Bartel, "MicroRNA regulation of gene expression in plants," *Current Opinion in Plant Biology*, vol. 7, no. 5, pp. 512–520, 2004.
- [12] F. R. Kulcheski, L. F. V. de Oliveira, L. G. Molina et al., "Identification of novel soybean microRNAs involved in abiotic and biotic stresses," *BMC Genomics*, vol. 12, article 307, 2011.
- [13] D. P. Bartel, "MicroRNAs: genomics, biogenesis, mechanism, and function," *Cell*, vol. 116, no. 2, pp. 281–297, 2004.
- [14] V. Ambros, R. C. Lee, A. Lavanway, P. T. Williams, and D. Jewell, "MicroRNAs and other tiny endogenous RNAs in *C. elegans*," *Current Biology*, vol. 13, no. 10, pp. 807–818, 2003.
- [15] G. Tang, B. J. Reinhart, D. P. Bartel, and P. D. Zamore, "A biochemical framework for RNA silencing in plants," *Genes & Development*, vol. 17, no. 1, pp. 49–63, 2003.
- [16] X. Chen, "microRNA biogenesis and function in plants," *FEBS Letters*, vol. 579, no. 26, pp. 5923–5931, 2005.
- [17] Y. Kurihara and Y. Watanabe, "Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions," *Proceedings of the*

- National Academy of Sciences of the United States of America*, vol. 101, no. 34, pp. 12753–12758, 2004.
- [18] V. N. Kim, “MicroRNA biogenesis: coordinated cropping and dicing,” *Nature Reviews Molecular Cell Biology*, vol. 6, no. 5, pp. 376–385, 2005.
- [19] P. Brodersen, L. Sakvarelidze-Achard, M. Bruun-Rasmussen et al., “Widespread translational inhibition by plant miRNAs and siRNAs,” *Science*, vol. 320, no. 5880, pp. 1185–1190, 2008.
- [20] L. Guo and Z. Lu, “Global expression analysis of miRNA gene cluster and family based on isomiRs from deep sequencing data,” *Computational Biology and Chemistry*, vol. 34, no. 3, pp. 165–171, 2010.
- [21] W. Park, J. Li, R. Song, J. Messing, and X. Chen, “CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*,” *Current Biology*, vol. 12, no. 17, pp. 1484–1495, 2002.
- [22] A. Kozomara and S. Griffiths-Jones, “MiRBase: annotating high confidence microRNAs using deep sequencing data,” *Nucleic Acids Research*, vol. 42, no. 1, pp. D68–D73, 2014.
- [23] H. Li, Y. Dong, H. Yin et al., “Characterization of the stress associated microRNAs in *Glycine max* by deep sequencing,” *BMC Plant Biology*, vol. 11, article 170, 2011.
- [24] Q.-H. Zhu, A. Spriggs, L. Matthew et al., “A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains,” *Genome Research*, vol. 18, no. 9, pp. 1456–1465, 2008.
- [25] W. Jin, N. Li, B. Zhang et al., “Identification and verification of microRNA in wheat (*Triticum aestivum*),” *Journal of Plant Research*, vol. 121, no. 3, pp. 351–355, 2008.
- [26] R. L. R. Pilcher, S. Moxon, N. Pakseresht et al., “Identification of novel small RNAs in tomato (*Solanum lycopersicum*),” *Planta*, vol. 226, no. 3, pp. 709–717, 2007.
- [27] D. Ding, L. Zhang, H. Wang, Z. Liu, Z. Zhang, and Y. Zheng, “Differential expression of miRNAs in response to salt stress in maize roots,” *Annals of Botany*, vol. 103, no. 1, pp. 29–38, 2009.
- [28] J. T. Cuperus, N. Fahlgren, and J. C. Carrington, “Evolution and functional diversification of *MIRNA* genes,” *The Plant Cell*, vol. 23, no. 2, pp. 431–442, 2011.
- [29] R. Rajagopalan, H. Vaucheret, J. Trejo, and D. P. Bartel, “A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*,” *Genes & Development*, vol. 20, no. 24, pp. 3407–3425, 2006.
- [30] J. Zhu, W. Li, W. Yang, L. Qi, and S. Han, “Identification of microRNAs in *Caragana intermedia* by high-throughput sequencing and expression analysis of 12 microRNAs and their targets under salt stress,” *Plant Cell Reports*, vol. 32, no. 9, pp. 1339–1349, 2013.
- [31] K. P. McCormick, M. R. Willmann, and B. C. Meyers, “Experimental design, preprocessing, normalization and differential expression analysis of small RNA sequencing experiments,” *Silence*, vol. 2, no. 1, article 2, 2011.
- [32] R. Sunkar, X. Zhou, Y. Zheng, W. Zhang, and J.-K. Zhu, “Identification of novel and candidate miRNAs in rice by high throughput sequencing,” *BMC Plant Biology*, vol. 8, article 25, 2008.
- [33] G. Szittyá, S. Moxon, D. M. Santos et al., “High-throughput sequencing of *Medicago truncatula* short RNAs identifies eight new miRNA families,” *BMC Genomics*, vol. 9, article 593, 2008.
- [34] V. Pantaleo, G. Szittyá, S. Moxon et al., “Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis,” *The Plant Journal*, vol. 62, no. 6, pp. 960–976, 2010.
- [35] S. Moxon, R. Jing, G. Szittyá et al., “Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening,” *Genome Research*, vol. 18, no. 10, pp. 1602–1609, 2008.
- [36] C. Song, C. Wang, C. Zhang et al., “Deep sequencing discovery of novel and conserved microRNAs in trifoliolate orange (*Citrus trifoliata*),” *BMC Genomics*, vol. 11, no. 1, article 431, 2010.
- [37] Q.-X. Song, Y.-F. Liu, X.-Y. Hu et al., “Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing,” *BMC Plant Biology*, vol. 11, article 5, 2011.
- [38] C.-Z. Zhao, H. Xia, T. P. Frazier et al., “Deep sequencing identifies novel and conserved microRNAs in peanuts (*Arachis hypogaea* L.),” *BMC Plant Biology*, vol. 10, article 3, 2010.
- [39] D. Klevebring, N. R. Street, N. Fahlgren et al., “Genome-wide profiling of *Populus* small RNAs,” *BMC Genomics*, vol. 10, article 620, 2009.
- [40] S. Paul, A. Kundu, and A. Pal, “Identification and expression profiling of *Vigna mungo* microRNAs from leaf small RNA transcriptome by deep sequencing,” *Journal of Integrative Plant Biology*, vol. 56, no. 1, pp. 15–23, 2014.
- [41] A. Roy, A. Bandyopadhyay, A. K. Mahapatra et al., “Evaluation of genetic diversity in jute (*Corchorus* species) using STMS, ISSR and RAPD markers,” *Plant Breeding*, vol. 125, no. 3, pp. 292–297, 2006.
- [42] M. K. Sinha, S. Mitra, T. Ramasubramanian, and B. S. Mahapatra, “Crop diversification for profitability in jute and allied fibre crops,” *Indian Journal of Agronomy*, vol. 54, no. 2, pp. 221–225, 2009.
- [43] S. Ahmed, M. D. Shafiuddin, M. S. Azam, M. S. Islam, A. Ghosh, and H. Khan, “Identification and characterization of jute LTR retrotransposons: their abundance, heterogeneity and transcriptional activity,” *Mobile Genetic Elements*, vol. 1, no. 1, pp. 18–28, 2011.
- [44] R. Samira, M. M. Moosa, M. M. Alam, S. I. Keka, and H. Khan, “‘In silico’ analysis of jute SSR library and experimental verification of assembly,” *Plant OMICS*, vol. 3, no. 2, pp. 57–65, 2010.
- [45] R. Schwab, J. F. Palatnik, M. Riester, C. Schommer, M. Schmid, and D. Weigel, “Specific effects of microRNAs on the plant transcriptome,” *Developmental Cell*, vol. 8, no. 4, pp. 517–527, 2005.
- [46] X. Dai and P. X. Zhao, “psRNATarget: a plant small RNA target analysis server,” *Nucleic Acids Research*, vol. 39, no. 2, pp. W155–W159, 2011.
- [47] M. Ashburner, C. A. Ball, J. A. Blake et al., “Gene ontology: tool for the unification of biology,” *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.
- [48] M. Kanehisa, M. Araki, S. Goto et al., “KEGG for linking genomes to life and the environment,” *Nucleic Acids Research*, vol. 36, no. 1, pp. D480–D484, 2008.
- [49] M. Kanehisa, S. Goto, M. Hattori et al., “From genomics to chemical genomics: new developments in KEGG,” *Nucleic Acids Research*, vol. 34, pp. D354–D357, 2006.
- [50] E. Varkonyi-Gasic, R. Wu, M. Wood, E. F. Walton, and R. P. Hellens, “Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs,” *Plant Methods*, vol. 3, no. 1, article 12, 2007.
- [51] C. Chen, D. A. Ridzon, A. J. Broomer et al., “Real-time quantification of microRNAs by stem-loop RT-PCR,” *Nucleic Acids Research*, vol. 33, no. 20, p. e179, 2005.

- [52] S. W. Burge, J. Daub, R. Eberhardt et al., "Rfam 11.0: 10 years of RNA families," *Nucleic Acids Research*, vol. 41, pp. D226–D232, 2012.
- [53] A. Kozomara and S. Griffiths-Jones, "miRBase: integrating microRNA annotation and deep-sequencing data," *Nucleic Acids Research*, vol. 39, no. 1, Article ID gkq1027, pp. D152–D157, 2011.
- [54] B. C. Meyers, M. J. Axtell, B. Bartel et al., "Criteria for annotation of plant microRNAs," *The Plant Cell*, vol. 20, no. 12, pp. 3186–3190, 2008.
- [55] K. Okamura, M. D. Phillips, D. M. Tyler, H. Duan, Y.-T. Chou, and E. C. Lai, "The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution," *Nature Structural & Molecular Biology*, vol. 15, no. 4, pp. 354–363, 2008.
- [56] S. Yang Jr., M. D. Phillips, D. Betel et al., "Widespread regulatory activity of vertebrate microRNA* species," *RNA*, vol. 17, no. 2, pp. 312–326, 2011.
- [57] P. Nelson, M. Kiriakidou, A. Sharma, E. Maniatakis, and Z. Mourelatos, "The microRNA world: small is mighty," *Trends in Biochemical Sciences*, vol. 28, no. 10, pp. 534–540, 2003.
- [58] M. Zuker, "Mfold web server for nucleic acid folding and hybridization prediction," *Nucleic Acids Research*, vol. 31, no. 13, pp. 3406–3415, 2003.
- [59] D. Li, L. Wang, X. Liu et al., "Deep sequencing of maize small RNAs reveals a diverse set of microRNA in dry and imbibed seeds," *PLoS ONE*, vol. 8, no. 1, Article ID e55107, 2013.
- [60] W. Mao, Z. Li, X. Xia, Y. Li, and J. Yu, "A combined approach of high-throughput sequencing and degradome analysis reveals tissue specific expression of microRNAs and their targets in cucumber," *PLoS ONE*, vol. 7, no. 3, Article ID e33040, 2012.
- [61] Q.-Y. Zeng, C.-Y. Yang, Q.-B. Ma, X.-P. Li, W.-W. Dong, and H. Nian, "Identification of wild soybean miRNAs and their target genes responsive to aluminum stress," *BMC Plant Biology*, vol. 12, article 182, 2012.
- [62] M. Y. Xu, Y. Dong, Q. X. Zhang et al., "Identification of miRNAs and their targets from Brassica napus by high-throughput sequencing and degradome analysis," *BMC Genomics*, vol. 13, no. 1, article 421, 2012.
- [63] N. Fahlgren, M. D. Howell, K. D. Kasschau et al., "High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of MIRNA genes," *PLoS ONE*, vol. 2, no. 2, article e219, 2007.
- [64] Y. Yao, G. Guo, Z. Ni et al., "Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.)," *Genome Biology*, vol. 8, no. 6, article R96, 2007.
- [65] X.-J. Wang, J. L. Reyes, N.-H. Chua, and T. Gaasterland, "Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets," *Genome Biology*, vol. 5, no. 9, p. R65, 2004.
- [66] C. Llave, K. D. Kasschau, M. A. Rector, and J. C. Carrington, "Endogenous and silencing-associated small RNAs in plants," *The Plant Cell*, vol. 14, no. 7, pp. 1605–1619, 2002.
- [67] D. M. O'Malley, R. Whetten, W. Bao, C.-L. Chen, and R. R. Sederoff, "The role of laccase in lignification," *The Plant Journal*, vol. 4, no. 5, pp. 751–757, 1993.
- [68] G. van Ooijen, G. Mayr, M. M. A. Kasiem, M. Albrecht, B. J. C. Cornelissen, and F. L. W. Takken, "Structure-function analysis of the NB-ARC domain of plant disease resistance proteins," *Journal of Experimental Botany*, vol. 59, no. 6, pp. 1383–1397, 2008.
- [69] Gene Ontology Consortium, "The Gene Ontology (GO) database and informatics resource," *Nucleic Acids Research*, vol. 32, pp. D258–D261, 2004.
- [70] E. Altermann and T. R. Klaenhammer, "PathwayVoyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database," *BMC Genomics*, vol. 6, article 60, 2005.
- [71] V. Benes and M. Castoldi, "Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available," *Methods*, vol. 50, no. 4, pp. 244–249, 2010.
- [72] K. Etebari and S. Asgari, "Accuracy of microRNA discovery pipelines in non-model organisms using closely related species genomes," *PLoS ONE*, vol. 9, no. 1, Article ID e84747, 2014.
- [73] J.-Z. Zhang, X.-Y. Ai, W.-W. Guo, S.-A. Peng, X.-X. Deng, and C.-G. Hu, "Identification of miRNAs and their target genes using deep sequencing and degradome analysis in trifoliolate orange [*Poncirus trifoliata* (L.) Raf.]," *Molecular Biotechnology*, vol. 51, no. 1, pp. 44–57, 2012.
- [74] M. W. Jones-Rhoades, D. P. Bartel, and B. Bartel, "MicroRNAs and their regulatory roles in plants," *Annual Review of Plant Biology*, vol. 57, pp. 19–53, 2006.
- [75] G. Wu and R. S. Poethig, "Temporal regulation of shoot development in *Arabidopsis thaliana* by miRr156 and its target SPL3," *Development*, vol. 133, no. 18, pp. 3539–3547, 2006.
- [76] M. J. Axtell and D. P. Bartel, "Antiquity of microRNAs and their targets in land plants," *The Plant Cell*, vol. 17, no. 6, pp. 1658–1673, 2005.
- [77] E. Allen, Z. Xie, A. M. Gustafson, G.-H. Sung, J. W. Spatafora, and J. C. Carrington, "Evolution of microRNA genes by inverted duplication of target gene sequences in *Arabidopsis thaliana*," *Nature Genetics*, vol. 36, no. 12, pp. 1282–1290, 2004.
- [78] M. J. Axtell and J. L. Bowman, "Evolution of plant microRNAs and their targets," *Trends in Plant Science*, vol. 13, no. 7, pp. 343–349, 2008.
- [79] G. Martínez, J. Forment, C. Llave, V. Pallás, and G. Gómez, "High-throughput sequencing, characterization and detection of new and conserved cucumber miRNAs," *PLoS ONE*, vol. 6, no. 5, Article ID e19523, 2011.
- [80] E. A. Glazov, P. A. Cottee, W. C. Barris, R. J. Moore, B. P. Dalrymple, and M. L. Tizard, "A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach," *Genome Research*, vol. 18, no. 6, pp. 957–964, 2008.
- [81] S.-J. Kou, X.-M. Wu, Z. Liu, Y.-L. Liu, Q. Xu, and W.-W. Guo, "Selection and validation of suitable reference genes for miRNA expression normalization by quantitative RT-PCR in citrus somatic embryogenic and adult tissues," *Plant Cell Reports*, vol. 31, no. 12, pp. 2151–2163, 2012.
- [82] Y. Ao, Y. Wang, L. Chen, T. Wang, H. Yu, and Z. Zhang, "Identification and comparative profiling of microRNAs in wild-type *Xanthoceras sorbifolia* and its double flower mutant," *Genes & Genomics*, vol. 34, no. 5, pp. 561–568, 2012.
- [83] F. Vazquez, T. Blevins, J. Ailhas, T. Boller, and F. Meins Jr., "Evolution of *Arabidopsis* MIR genes generates novel microRNA classes," *Nucleic Acids Research*, vol. 36, no. 20, pp. 6429–6438, 2008.
- [84] A. Git, H. Dvinge, M. Salmon-Divon et al., "Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression," *RNA*, vol. 16, no. 5, pp. 991–1006, 2010.
- [85] S. Tam, R. de Borja, M.-S. Tsao, and J. D. McPherson, "Robust global microRNA expression profiling using next-generation

- sequencing technologies,” *Laboratory Investigation*, vol. 94, no. 3, pp. 350–358, 2014.
- [86] K. Xie, C. Wu, and L. Xiong, “Genomic organization, differential expression, and interaction of SQUAMOSA promoter-binding-like transcription factors and microRNA156 in rice,” *Plant Physiology*, vol. 142, no. 1, pp. 280–293, 2006.
- [87] H. Zhu, F. Hu, R. Wang et al., “*Arabidopsis* argonaut10 specifically sequesters miR166/165 to regulate shoot apical meristem development,” *Cell*, vol. 145, no. 2, pp. 242–256, 2011.
- [88] M.-F. Wu, Q. Tian, and J. W. Reed, “*Arabidopsis* microRNA 167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction,” *Development*, vol. 133, no. 21, pp. 4211–4218, 2006.
- [89] H. K. Jin, R. W. Hye, J. Kim et al., “Trifurcate feed-forward regulation of age-dependent cell death involving miR164 in *Arabidopsis*,” *Science*, vol. 323, no. 5917, pp. 1053–1057, 2009.
- [90] M. J. Aukerman and H. Sakai, “Regulation of flowering time and floral organ identity by a microRNA and its *APETALA2*-like target genes,” *The Plant Cell*, vol. 15, no. 11, pp. 2730–2741, 2003.
- [91] J. L. Reyes and N.-H. Chua, “ABA induction of miR159 controls transcript levels of two MYB factors during *Arabidopsis* seed germination,” *The Plant Journal*, vol. 49, no. 4, pp. 592–606, 2007.
- [92] B. Zhang, X. Pan, G. P. Cobb, and T. A. Anderson, “Plant microRNA: a small regulatory molecule with big impact,” *Developmental Biology*, vol. 289, no. 1, pp. 3–16, 2006.
- [93] J. B. Song, S. Q. Huang, T. Dalmay, and Z. M. Yang, “Regulation of leaf morphology by MicroRNA394 and its target *LEAF CURLING RESPONSIVENESS*,” *Plant and Cell Physiology*, vol. 53, no. 7, pp. 1283–1294, 2012.
- [94] C. G. Kawashima, C. A. Matthewman, S. Huang et al., “Interplay of SLIM1 and miR395 in the regulation of sulfate assimilation in *Arabidopsis*,” *Plant Journal*, vol. 66, no. 5, pp. 863–876, 2011.
- [95] D. V. Dugas and B. Bartel, “Sucrose induction of *Arabidopsis* miR398 represses two Cu/Zn superoxide dismutases,” *Plant Molecular Biology*, vol. 67, no. 4, pp. 403–417, 2008.
- [96] K. Aung, S.-I. Lin, C.-C. Wu, Y.-T. Huang, C.-L. Su, and T.-J. Chiou, “pho2, a phosphate overaccumulator, is caused by a nonsense mutation in a microRNA399 target gene,” *Plant Physiology*, vol. 141, no. 3, pp. 1000–1011, 2006.
- [97] Z. Xie, K. D. Kasschau, and J. C. Carrington, “Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation,” *Current Biology*, vol. 13, no. 9, pp. 784–789, 2003.
- [98] P. Ranocha, M. Chabannes, S. Chamayou et al., “Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar,” *Plant Physiology*, vol. 129, no. 1, pp. 145–155, 2002.
- [99] A. Driouich, A.-C. Lainé, B. Vian, and L. Paye, “Characterization and localization of laccase forms in stem and cell cultures of sycamore,” *The Plant Journal*, vol. 2, no. 1, pp. 13–24, 1992.
- [100] R. Sterjiades, J. F. D. Dean, and K.-E. L. Eriksson, “Laccase from sycamore maple (*Acer pseudoplatanus*) polymerizes monolignols,” *Plant Physiology*, vol. 99, no. 3, pp. 1162–1168, 1992.
- [101] W. Bao, D. M. O’malley, R. Whetten, and R. R. Sederoff, “A laccase associated with lignification in loblolly pine xylem,” *Science*, vol. 260, no. 5108, pp. 672–674, 1993.
- [102] G. Sengupta and P. Palit, “Characterization of a lignified secondary phloem fibre-deficient mutant of jute (*Corchorus capsularis*),” *Annals of Botany*, vol. 93, no. 2, pp. 211–220, 2004.
- [103] F. Chen and R. A. Dixon, “Lignin modification improves fermentable sugar yields for biofuel production,” *Nature Biotechnology*, vol. 25, no. 7, pp. 759–761, 2007.
- [104] S. Lu, Q. Li, H. Wei et al., “Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in *Populus trichocarpa*,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 26, pp. 10848–10853, 2013.