

## Research Article

# Discovery of Novel Leaf Rust Responsive microRNAs in Wheat and Prediction of Their Target Genes

Dhananjay Kumar,<sup>1</sup> Dharmendra Singh,<sup>1,2</sup> Pulkit Kanodia,<sup>1</sup> Kumble Vinod Prabhu,<sup>3</sup> Manish Kumar,<sup>1</sup> and Kunal Mukhopadhyay<sup>1</sup>

<sup>1</sup> Department of Biotechnology, Birla Institute of Technology, Mesra, Ranchi 835215, India

<sup>2</sup> Department of Biotechnology, Eternal University, Baru Sahib, Himachal Pradesh 173101, India

<sup>3</sup> National Phytotron Facility, IARI, New Delhi 110012, India

Correspondence should be addressed to Kunal Mukhopadhyay; [kmukhopadhyay@bitmesra.ac.in](mailto:kmukhopadhyay@bitmesra.ac.in)

Received 21 May 2014; Accepted 19 July 2014; Published 12 August 2014

Academic Editor: Ben Berkhout

Copyright © 2014 Dhananjay Kumar 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 are endogenous small noncoding RNAs which play critical roles in gene regulation. Few wheat (*Triticum aestivum* L.) miRNA sequences are available in miRBase repertoire and knowledge of their biological functions related to biotic stress is limited. We identified 52 miRNAs, belonging to 19 families, from next-generation transcriptome sequence data based on homology search. One wheat specific novel miRNA was identified but could not be ascribed or assigned to any known miRNA family. Differentially expressed 22 miRNAs were found between susceptible and resistant wheat near-isogenic lines inoculated with leaf rust pathogen *Puccinia triticina* and compared with mock inoculated controls. Most miRNAs were more upregulated in susceptible NIL compared to resistant NIL. We identified 1306 potential target genes for these 52 miRNAs with vital roles in response to stimuli, signaling, and diverse metabolic and cellular processes. Gene ontology analysis showed 66, 20, and 35 target genes to be categorized into biological process, molecular function, and cellular component, respectively. A miRNA-mediated regulatory network revealed relationships among the components of the targetome. The present study provides insight into potential miRNAs with probable roles in leaf rust pathogenesis and their target genes in wheat which establish a foundation for future studies.

## 1. Introduction

Bread wheat (*Triticum aestivum* L.) is the second most important food crop after rice based on international commodity price and third in order of production after maize and rice. Wheat is extensively grown throughout the world and provides protein content, as well as 20% of basic caloric value to humans [1]. Wheat is the last major cereal crop for which complete genome sequence is not available due to its large genome size (16.9 GB), high proportion of transposable and repetitive DNA element, and presence of three closely related genomes (AABBDD,  $2n = 6x = 42$ ) [2]. However, recent technological advances such as next-generation sequencing platforms now offer large scale programs that delivered 5x genomic sequence resources for wheat [3]. Leaf rust is a major disease of wheat caused by the obligate biotrophic fungus *Puccinia triticina* Erikss that results in 10% yield loss

annually [4]. More than 60 leaf rust resistant (*Lr*) genes had been identified from wheat and its wild relatives that confer resistance in a gene-for-gene manner. Many of them had been introgressed into wheat, but due to continuous evolution of the pathogen genome resistance breaks down quickly [5]. The molecular basis underlying leaf rust-wheat interaction is still not completely understood and as of date is a major focus of research.

Small RNAs, broadly divided into microRNAs (miRNA) and small interfering RNAs (siRNA), are global regulators of gene expression mainly through posttranscriptional or occasionally by translational repression in plants [6–8]. They are noncoding, single-stranded RNAs of 18–24 nt in length and are characterized by stem-loop structures of their precursors [9]. MiRNA genes are located mostly within noncoding regions of genomes and usually transcribed from promoters by RNA polymerase II [10, 11]. MiRNA genes

represent about 1-2% of the known eukaryotic genomes and constitute an important class of fine-tuning regulators that reprogram numerous transcription events implicated in several physiological or disease-associated cellular processes by interacting with each other in regulatory networks [12, 13]. Studies in plants have revealed their key roles in diverse regulatory pathways, including auxin signaling, meristem boundary formation, organ separation, leaf development and polarity, lateral root formation, transition from juvenile to adult vegetative phase and from vegetative to flowering phase, floral organ identity and reproduction, and defense response against different biotic and abiotic stress [14, 15]. A total of 153 miRNAs were identified in wheat, of which 24 and 12 miRNAs were responsive to powdery mildew infection and heat stress, respectively [16]. The authors using *in silico* techniques also predicted 149 target genes which are potentially regulated by those miRNAs. In rice, a “osa-miR7695” miRNA through natural selection and domestication events during evolution negatively regulates an alternatively spliced transcript of natural resistance-associated macrophage protein 6 (OsNramp6) [14]. Overexpression of the miRNA in rice also conferred blast fungal pathogen resistance. This study highlights a miRNA-mediated regulation of OsNramp6 in disease resistance, whilst illustrating the existence of a novel regulatory network that integrates miRNA function and mRNA processing in relation to plant immunity.

Considering the importance of miRNAs in gene regulation, two major categories of approaches have been applied for their investigation [17]. Besides the labor-intensive experimental approaches, comparative genomics-based computational strategies are faster, affordable, and effective and are the major contributors to miRBase repertoire [17]. Since most miRNAs are evolutionary and highly conserved, comparative genomics-based prediction of new miRNA homologous from previously known miRNA sequences is a powerful approach [18]. Different computational miRNA finding strategies have been developed based on the core principle of searching for conserved sequences between different species that can fold into extended hairpins [19]. The biogenesis of miRNAs suggests that it is possible to mine miRNAs and elucidate their functions by searching repository of available expressed sequence tags (ESTs) with known miRNAs in absence of a sequenced genome [18, 20–22]. Currently, several computational approaches have been used to successfully identify potential miRNA targets in mRNA sequences or to select potential targets for experimental validation [23].

Next-generation sequencing technologies provide a rapid and high-throughput approach to enable discovery of novel species-specific or low-abundance miRNAs. Since its first use in the model species *Arabidopsis* [24], high-throughput sequencing technologies have been successfully applied in many plant species, such as rice [25], poplar [26], grapevine [27], *Medicago truncatula* [28], *Taxus chinensis* [29], foxtail millet [30], wheat, and *Brachypodium* [31] for identification of miRNAs. A total of 58 wheat miRNAs were identified by 454 sequencing, [32] and 81 novel miRNAs were identified through Solexa high-throughput sequencing [16]. In a recent study, numerous miRNAs were predicted on chromosome 6B of wheat; many of them located on repeat regions, DNA

transposons, and retrotransposons suggesting propagation of miRNA genes in the wheat genome [33]. Although wheat is one of the most cultivated crops in the world, limited researches have been performed on wheat miRNAs and no studies have been reported on roles of miRNAs during wheat-leaf rust infection. In this study, new miRNAs were mined from the next-generation transcriptome sequencing data prepared from mock and pathogen inoculated wheat near-isogenic lines (NILs) for the purpose of understanding their roles in rust disease development, metabolism, and other associated physiological processes.

## 2. Materials and Methods

**2.1. Plant Material, Pathogen, and Inoculation Treatments.** Two wheat lines, HD2329, a seedling leaf rust susceptible phenotype, and its NIL HD2329+*Lr28*, a seedling leaf rust resistant with nest-immune (0-0;) phenotype, were selected for the study. The *Lr28* gene was derived from *Aegilops speltoides* (Tausch) that is effective against all pathotypes of the pathogen in India [34]. *Puccinia triticina* pathotype 77-5, the most predominant and devastating pathotype in all parts of the Indian subcontinent, was selected as experimental pathogen. Plant growth, pathogen inoculum preparation, inoculation of the pathogen, and infectivity screening were performed at the National Phytotron Facility, IARI, New Delhi, exactly as mentioned in Singh et al. [35].

**2.2. RNA Isolation, SAGE Library Preparation, and SOLiD.** Leaf tissues were pooled from 15 seedlings; each mock and pathogen inoculated NILs at 24 hours after inoculation (hpi) and quickly dipped in liquid nitrogen. Total RNA was isolated using TRI Reagent (Molecular Research Center Inc., USA) as recommended by manufacturer. The RNA isolation time-point was based on earlier studies on development of infection structures [36] and activation of resistant signaling genes [35, 37]. The integrity of the isolated RNAs was confirmed on an Agilent Bioanalyser 2100. Four serial analysis of gene expression (SAGE) libraries were prepared from the isolated RNAs [coded as (i) S-M: HD2329 mock inoculated, (ii) S-PI: HD2329 pathogen inoculated, (iii) R-M: HD2329+*Lr28* mock inoculated, and (iv) R-PI: HD2329+*Lr28* pathogen inoculated] using SOLiD SAGE kit (Applied Biosystems, CA, USA) following the recommended protocol and sequenced using Sequencing by Oligonucleotide Ligation and Detection (SOLiD) technique commercially. The sequences had been submitted to NCBI SRA061917 (BioSample accession as SAMM01820702, SAMM01820703, SAMM01820704, and SAMM01820705).

**2.3. Prediction of Novel and Conserved miRNA in Wheat.** After removing adaptor and redundant and ambiguous sequences, readings smaller than <18 and >28 nucleotides (nt) were discarded using small RNA analysis tool (CLC genomics workbench) [38]. Novel and conserved miRNAs were identified in wheat by annotating SAGE tags from the four libraries (S-M, S-PI, R-M, and R-PI) to existing miRNA sequences of *Triticum aestivum*, *Triticum turgidum*, *Aegilops tauschii*, *Brachypodium distachyon*, *Hordeum vulgare*, *Zea*

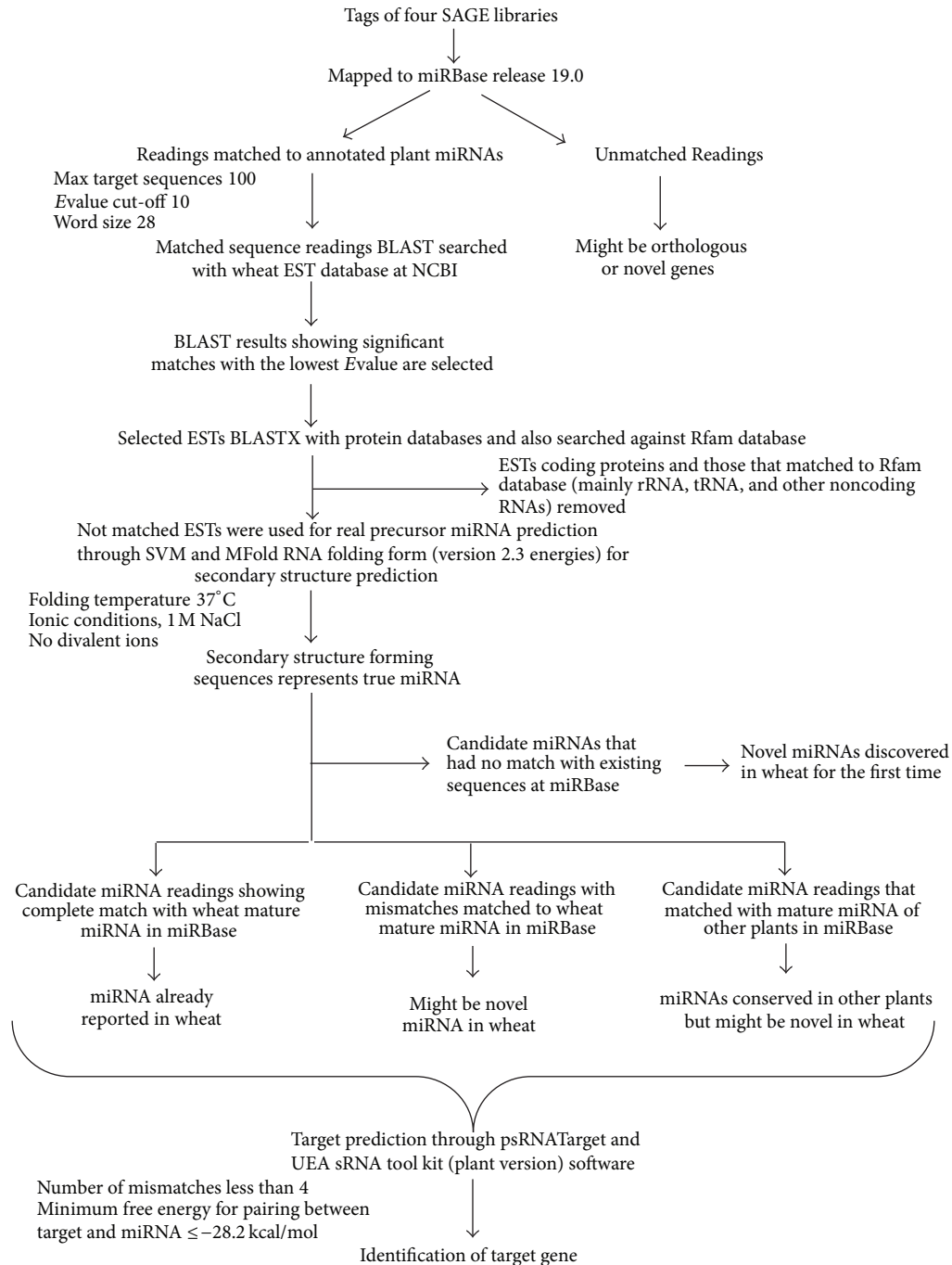


FIGURE 1: Schematic presentation of the pipeline developed to discover miRNA from NGS data and target prediction of miRNAs. Sequences of miRNAs for *Triticum aestivum*, *Triticum turgidum*, *Aegilops tauschii*, *Brachypodium distachyon*, *Hordeum vulgare*, *Zea mays*, *Oryza sativa*, *Saccharum* ssp., *Sorghum bicolor*, *Brassica* ssp., *Arabidopsis thaliana*, *Glycine max*, *Vitis vinifera*, and *Solanum lycopersicum* were used for homology searches in miRBase.

*mays*, *Oryza sativa*, *Saccharum* ssp., *Sorghum bicolor*, *Brassica* ssp., *Arabidopsis thaliana*, *Glycine max*, *Vitis vinifera*, and *Solanum lycopersicum* available at miRBase (release 19, August 2012) [39, 40]. A pipeline mentioning the strategies and software used to identify and confirm the miRNAs was developed following Yao et al. [41] and Sunkar et al. [25] (Figure 1). The number of times each miRNA is represented

in a particular SAGE library served as an index for estimation of their relative abundance (counts).

The initial BLASTN search with annotated miRNAs to wheat ESTs was carried out. To increase the potential of identification of miRNAs, more stringent parameters were used as follows: (1) the default word-match size was set at seven, the smallest setting that can be used for the online BLASTN

TABLE 1: Statistics of small RNA sequences in all four libraries.

| Library | Tag counts after removing low quality sequences | Average length of tags (nt) | Tags annotated to miRBase | Number of ESTs matched to annotated tags | Number of ESTs after BLASTX and Rfam searches | Real precursor miRNAs conformed by SVM |
|---------|---|-----------------------------|---------------------------|--|---|--|
| S-M     | 4,949,795                                       | 29.5                        | 258                       | 3253                                     | 1178  | 42                                     |
| S-PI    | 4,712,304                                       | 29.3                        | 432                       | 4750                                     | 2293  | 43                                     |
| R-M     | 3,384,144                                       | 28.6                        | 185                       | 1730                                     | 694   | 30                                     |
| R-PI    | 3,021,557                                       | 28.4                        | 368                       | 4021                                     | 1298  | 27                                     |

program between the miRNA query and the EST sequences; (2) the expected values were set at 1,000 to increase the hit chance for more potential sequences; and (3) the sequence number of the BLAST search and the sequence alignments was set to 1,000. MiRNA sequences matching at least 18 nt and mismatches less than 3 nt were selected for further analysis. The protein coding ESTs sequences were removed with the help of BLASTX program; and only noncoding sequences were retained. BLASTN search was performed against Rfam 11.0 (<http://rfam.sanger.ac.uk/>) to distinguish between miRNA and other small RNA families such as rRNA, snRNA, and tRNA. Nonredundant ESTs sequences were selected with the help of Perl program. Real miRNA precursors were predicted through triplet-SVM classifier program based on support vector machine [42]. This software needs additional packages like RNAfold and LibSVM. The secondary structure of pre-miRNAs was determined using MFOLD 3.2 software [43]. The following criteria were considered for screening the candidate miRNA homologs to significantly reduce false positives as proposed by Ambros et al. [44] and Meyers et al. [32]: (i) the RNA sequence folding into an appropriate stem-loop hairpin secondary structure should contain the ~18 nt mature miRNA sequence located on one arm of the hairpin structure, (ii) the predicted mature miRNAs should not have more than 6 mismatches with the miRNA\* sequence on the opposite arm, (iii) maximum size of 3 nt was allowed for a bulge on the miRNA sequence, (iv) miRNA precursors with secondary structures should have low minimal free energy  $-18$  Kcal mole<sup>-1</sup> and minimal free energy index (MFEI) [45], (v) the A + U content of pre-miRNA within the range of 30–70% was considered, and (vi) no loop or break in miRNA sequences was allowed. The MFEI was calculated using the equation  $MFEI = [(MFE/\text{length of the pre-miRNA sequence}) \times 100]/(G + C) \%$  where MFE denotes the negative folding free energies ( $\Delta G$ ). The unique miRNA sequence that perfectly matched with wheat miRNAs in miRBase was considered as known miRNAs, whereas those with partial match might be novel miRNAs for wheat. On the other hand, miRNA sequences that matched with mature miRNA sequence of other plants might be conserved in those plants but could be novel in wheat. The sequences that had no match with existing wheat miRNA sequences at miRBase were considered novel for wheat [46]. We also compared the miRNAs obtained in this study with those previously reported by Yao et al. [41] and Han et al. [20].

**2.4. Prediction of miRNA Targets Genes and Their Functions.** The putative target sites of identified miRNAs were

predicted using two mostly used plant miRNA-target prediction algorithms, plant target prediction tool available on psRNATarget [47] and UEA sRNA tool kit [48] (Figure 1). Binding of miRNAs to the targets with perfect or nearly perfect complementarities was considered as this influences transcript regulation. Gaps and more than 4 mismatches between mature miRNAs and their potential target mRNA were not accepted. We subjected the potential miRNA-target genes to a functional enrichment analysis using Blast2GO software (v2.5.1) [49] with the default parameters used to obtain the GO terms for each target wheat gene. The WEGO online tool [50] was used to perform a GO enrichment analysis of the miRNA targets.

### 3. Results

**3.1. Identification and Characterization of Potential Wheat miRNAs from SAGE Libraries.** After filtering the low quality sequences, adapter sequences, and ambiguous nt, the number of tags in S-M, S-PI, R-M, and R-PI libraries was 4,949,795, 4,712,304, 3,384,144, and 3,021,557, respectively, and the average length of the sequences was 28.4–29.5 nt (Table 1). This final dataset of high quality readings was used to identify the miRNAs. The initial analysis following the developed bioinformatics pipeline (Figure 1) uncovered 258, 432, 185, and 368 predictive miRNA sequences having tag counts of 6521, 49296, 2563, and 6384 in S-M, S-PI, R-M, and R-PI libraries, respectively. The large number of miRNAs identified in the present study in wheat suggests the wide existence of miRNAs in monocot species that might play important roles in biotic stress besides growth and development. Alteration in tag counts also indicated differential expression of the identified miRNAs in response to leaf rust pathogenesis. We obtained 42, 43, 30, and 27 miRNAs from S-M, S-PI, R-M, and R-PI libraries, respectively, that provided a total of 52 unique miRNAs. All these predicted 52 miRNAs successfully accreted the criteria for identification of candidate miRNA [44]. Many miRNAs could not be uncovered due to availability of limited EST sequences and lack of completely sequenced wheat genome.

The identified mature miRNA lengths varied from 20 to 25 nt (average  $21.21 \pm 1.47$  nt) with the 20 nt length being the predominant size (Figure 2(a)) indicating the typical size range for Dicer-derived products [51]. The pre-miRNA lengths ranged from 59 to 660 nt, with an average length of  $139.76 \pm 91.76$  nt (Figure 2(b); Table 2). However, a large percentage of the pre-miRNAs (51.92%) were 80–130 nt and only 6 pre-miRNAs (11.53%) were more than 230 nt in length.

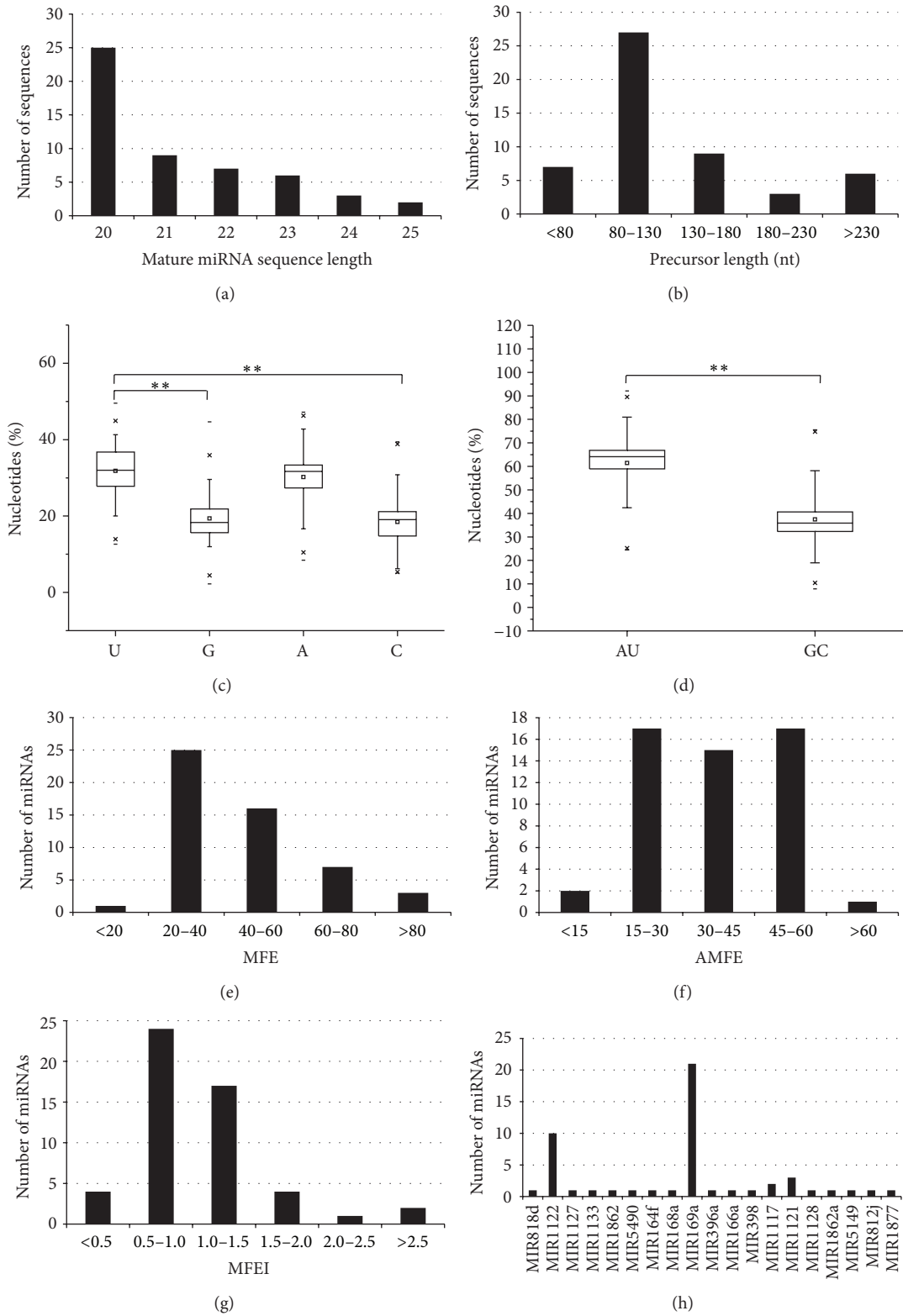


FIGURE 2: Characterization of miRNAs identified in wheat. Length distribution of mature miRNAs (a) and pre-miRNAs (b); the distribution of four nucleotides (c and d), MFE (e), AMFE (f), and MFEI (g) in the pre-miRNA sequences and miRNA families (h).



TABLE 2: Statistics of the characterized parameters of wheat miRNA precursors.

| Parameter                      | Mean   | Standard deviation | Minimum | Maximum |
|--------------------------------|--------|--------------------|---------|---------|
| MFE ( $\Delta G$ , -kcal/mol)  | 44.62  | 21.86              | 17.4    | 143.4   |
| AMFE ( $\Delta G$ , -kcal/mol) | 36.01  | 13.66              | 13.63   | 62.67   |
| MFEI ( $\Delta G$ , -kcal/mol) | 1.07   | 0.63               | 0.38    | 3.8     |
| Length (nt)                    | 139.76 | 91.76              | 59      | 660     |
| (G + C)%                       | 37.39  | 13.47              | 7.9     | 75.58   |
| (U + A)%                       | 61.4   | 13.56              | 24.42   | 92.13   |
| A%                             | 30.2   | 7.8                | 8.45    | 47.19   |
| C%                             | 18.46  | 7.03               | 5.52    | 39.43   |
| G%                             | 19.38  | 7.2                | 2.22    | 44.65   |
| U%                             | 31.83  | 7.13               | 12.62   | 49.65   |
| A/U ratio                      | 0.96   | 0.22               | 0.46    | 1.07    |
| C/G ratio                      | 1      | 0.36               | 0.41    | 2.5     |

The composition of the four nucleotides (A, G, C, and U) in pre-miRNA is an important indicator for species evolution and for the stabilization of RNA secondary structure. The percentage composition of each nucleotide was not evenly distributed in the identified wheat pre-miRNAs (Table 2). As in many previous studies the nucleotide uracil (U) is the most dominant in both mature miRNAs and pre-miRNAs in both plants and animals [52]. In this study also, we observed that the U content varied from 12.62% to 49.65% with an average of  $31.83 \pm 7.13\%$  in the identified wheat pre-miRNAs (Figure 2(c)). A majority (88.46%) of pre-miRNAs contained more than 25% of the nucleotide U (Figure 2(c)), which is significantly higher than C ( $18.46 \pm 7.03\%$ ) and G ( $19.38 \pm 7.20\%$ ) nucleotides (*t*-test,  $P < 1e - 5$ ). However, in the identified wheat pre-miRNAs, the GC content ( $37.39 \pm 13.47\%$ ) was significantly lower than the AU content ( $61.40 \pm 13.56\%$ ) (*t*-test,  $P < 1e - 5$ ) (Table 2; Figure 2(d)). Our results were consistent with previous findings on Asiatic cotton [52] and rainbow trout [53] where more than half of the total number of nucleotides was A or U. In this study, we also determined the A/U and C/G ratio to be 0.96 and 1.00, respectively, which suggest the presence of more U and C nucleotides in precursor sequences. This criterion also differentiates miRNAs from other types of cellular RNAs such as tRNAs, rRNAs, and mRNAs [54].

The major criterion for determining the stability of a nucleic acid secondary structure is the minimal folding free energy (MFE). Usually, the lower the MFE is, the more stable the secondary structure of the RNA molecule is. The MFE of the identified wheat pre-miRNAs varied from  $-17.40$  to  $-143.40$  kcal mol<sup>-1</sup> with an average of  $-44.62 \pm 21.86$  kcal mol<sup>-1</sup> (Table 2; Figure 2(e)). The MFEs for a majority of wheat pre-miRNAs ranged from  $-20$  to  $-40$  kcal mol<sup>-1</sup>. This large variation in MFE values signifies variation in their length. For better measurement of stability of RNA secondary structures, the adjusted minimal folding free energy (AMFE) strategy was used [55]. AMFE is the MFE of a RNA/DNA sequence that is 100 nt in length. The AMFE of the identified wheat pre-miRNAs ranged from  $-13.63$  to  $-62.67$  kcal mol<sup>-1</sup> with an average of  $-36.01 \pm 13.66$  kcal mol<sup>-1</sup> (Table 2), which

is a smaller range as compared with the MFE range. Most of the wheat pre-miRNAs identified in this study have AMFE values between  $-15$  and  $-60$  kcal mol<sup>-1</sup> (Figure 2(f)). The minimal folding free energy index (MFEI), a new standard for assaying miRNAs that can distinguish miRNAs from other coding and noncoding RNAs, revealed MFEI of the identified wheat pre-miRNAs from 0.38 to 3.8 with an average of  $1.07 \pm 0.63$  (Table 2; Figure 2(g)).

Using the BLASTN results and the secondary structure of the sequences, we identified 52 potential miRNAs from the high-throughput sequences of the SAGE libraries. Of these 52 miRNAs, 51 belonged to 19 families. However, the miRNAs identified in the present study were not evenly distributed among the families. The highest representation of 21 miRNAs in family MIR169a and 10 miRNAs in family MIR1122 was observed (Figure 2(h); see Supplementary Table S1 in the Supplementary Material available online at <http://dx.doi.org/10.1155/2014/570176>). MiRNAs belonging to the same family are derived from common ancestors and generally perform similar physiological processes. The function of MIR169a and MIR1122 families, having higher representation in the present study, contained miRNAs with defense responses to biotic and abiotic stresses. One wheat specific miRNA, TamiR18, was discovered from the dataset: but this miRNA could not be ascertained to any specific miRNA family. All other identified miRNAs have partial sequence similarity to known miRNAs from other plants but were identified in wheat for the first time (Supplementary Table S1). Therefore, these 51 miRNAs are conserved for other plants but identified for the first time in wheat. The majority (47 miRNAs) of the identified miRNAs were obtained from the plus strand. However, there were five miRNAs identified from the minus strand of certain ESTs (Supplementary Table S1). All predicted miRNA secondary structures showed 11–20 nt engaged in Watson-Crick or G/U base pairing between the mature miRNA and miRNA\* located on opposite arms of the hairpin structure (Supplementary Figure S1). The mature miRNA sequences could be located within either the 3' or the 5' arm of the secondary stem-loop hairpin structures (Supplementary Figure S1). Among the 52 identified miRNAs, 25



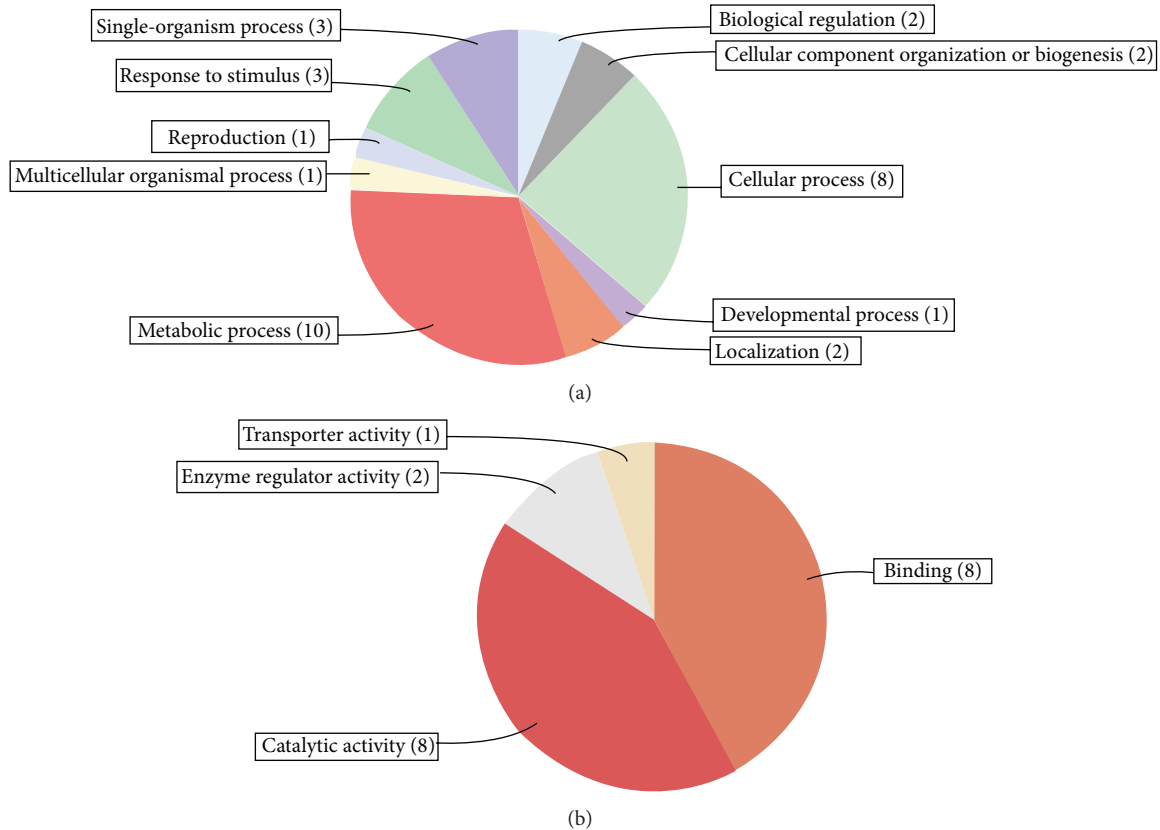


FIGURE 5: GO (gene ontology) term enrichment analysis of miRNAs target genes. Analysis of miRNAs targets within biological process (a) and molecular function (b) category.

For comprehensive annotation, all putative target transcripts were subjected to functional enrichment by gene ontology (GO) terms with the aid of Blast2GO and WEGO programs using default parameters. The identified wheat miRNA targets generated significant GO terms for further analysis. GO analysis categorized 66, 20, and 35 target genes into biological process, molecular function, and cellular component, respectively. The GO enrichment analysis showed that the predicted targets of the miRNAs were involved in a wide range of regulatory functions as well as some specific biological processes like metabolism, biosynthesis, and gene expression/transcription (Figure 5(a); Supplementary Figure S2). The transcripts representing genes with known functions were categorized by biological process and molecular function according to the ontological definitions of the GO terms. The majority of the predicted targets are involved in a broad range of biological processes such as metabolic process (10 terms), cellular process (8 terms), response to stimulus (3 terms), single-organism process (3 terms), developmental process (1 term), cellular component organization or biogenesis (2 terms), localization (2 terms), developmental process (1 term), reproduction (1 term), and biological regulation (2 terms) (Figure 5(a); Supplementary Figure S2). The putative target transcripts of miRNAs in the molecular function category were related to binding (8 terms), catalytic activity (8 terms), transporter activity (1 term), and enzyme regulator

activity (2 terms) (Figure 5(b); Supplementary Figure S2). Most of the miRNA-target genes assigned to the binding category were involved in nucleic acid binding, protein binding, and ion binding. Since these sequences are encoded for transcription factors, this is in accordance with previous reports [58–60]. In addition, some of the predicted targets had annotations in the cellular component category including cell and cell parts (Supplementary Figure S2).

Detailed analysis of the predicted targets revealed that the majority of the miRNAs have more than one potential regulatory target and conversely a single target could be regulated by more than one miRNA (Supplementary Table S2). This observation leads us to focus on development of a regulatory network comprising of more than one miRNA having different targets [46]. Based on the obtained miRNA-target pairs, we constructed a molecular network using Blast2GO software (Supplementary Figure S3). The constructed complex network shows significant relationship of the targetome with GO terms related to defense response, gene expression, metabolic process, cellular biosynthetic process, response to chemical stimulus, binding (ion binding, metal ion binding, RNA binding, DNA binding, ATP binding, etc.), and catalytic activity (transferase activity and hydrolase activity) (Supplementary Figures S3A and S3B). We were also able to identify seven wheat miRNAs having target genes involved in defense and signaling in plants (Table 3).



TABLE 3: Identified miRNAs having targets of pathogen secretome.

| Serial number | miRNA               | Target genes                   | Functions   |
|---------------|---------------------|--------------------------------|---|
| 1             | TamiR39             | Beta-glucanase                 | Have roles in plant defense by degrading cell walls of pathogen, thereby disrupting its deposition and contributing to pathogen death; subsequently, the released cell wall fragments act as elicitors for host defense response. |
| 2             | TamiR37,<br>TamiR38 | Peroxidase 6                   | Play important role in physiological processes like responses to biotic and abiotic stresses and biosynthesis of lignin and are involved in the scavenging of oxidative damage causing ROS.                                       |
| 3             | TamiR10,<br>TamiR43 | Calreticulin like protein      | Ubiquitous protein crucial for plant growth and development; powerful regulator in stress responses such as cold, drought, phytohormones, and pathogen.   |
| 4             | TamiR40             | Cinnamyl alcohol dehydrogenase | A key enzyme in lignin biosynthesis that catalyzes the final step in conversion of monolignols and also increases the strength and lodging resistance of stem tissues in monocot.   |
| 5             | TamiR36             | Pr1                            | Involved in plant defence responses against fungal pathogens  |

#### 4. Discussion

Numerous studies in recent years helped to discover and identify miRNAs from different plant species using next-generation sequencing approaches, followed by computational prediction methods [18, 19, 61–63]. Although numerous miRNAs have been identified in model plants like *Arabidopsis* and rice using experimental approaches, computational prediction from the EST sequences is the rapid and effective way to explore potential miRNAs in important crop plants like wheat where complete genomic sequence is still not available [17]. In the present study using comparative genomics-based search and bioinformatic tools we could identify 52 mature miRNAs in wheat belonging to 19 families. However, in addition to the highly conserved miRNAs, there are species-specific miRNAs originating from recently evolved *miRNA* genes [21, 64, 65]. These species-specific miRNAs often accumulate at lower levels and can be difficult to detect with traditional experimental based methods. The miRBase repository, as of release 19, has only 42 wheat miRNAs, a significantly inadequate number compared to other cereals (rice 708; maize 321, and *Sorghum* 242). This suggests that many more miRNAs are yet to be identified in wheat. Although the number of deeply conserved miRNA families in wheat largely remains the same as in *Arabidopsis*, all of the newly identified wheat miRNAs do not appear to be conserved in *Arabidopsis* and have predicted target genes with more diverse functions than those of conserved miRNAs.

Current literature suggests that several plant genes are involved in response to biotic and abiotic stress conditions exhibiting tissue and development stage specific expression which may be regulated at the posttranscriptional level by miRNA [41, 66, 67]. A large number of drought responsive miRNAs were identified in *Prunus persica* through high-throughput sequencing and their expression patterns were analyzed in roots and leaves [68]. At a genome-wide level 215 miRNAs and 447 miRNA targets were identified in foxtail millet and some novel foxtail millet miRNAs and their targets were validated experimentally [30]. Various pine-specific miRNA families were differentially expressed in response to infection by the rust fungus *Cronartium quercuum* f. sp. *fusiforme*, which causes fusiform rust disease

in pines [69]. Some of the conserved miRNAs differentially expressed in response to powdery mildew infection in wheat such as miR156, miR159, miR164, miR171, and miR396 were downregulated whereas miR393, miR444, and miR827 were upregulated, respectively [16].

The length of majority of mature miRNAs was 20 nt as was also found in Asiatic cotton [52]. A notable aspect of this study was accretion of one novel wheat specific miRNA TamiR18. We also compared the 52 TamiRNAs identified in the present study with the 58 and 37 TamiRNAs described by Yao et al. [41] and Han et al. [20]. The search did not yield any perfectly matched miRNAs. The wheat genotypes used in the present study and those used by Yao et al. [41] were different whereas Han et al. [20] used a different pipeline targeting the miRNA registry database available at the Sanger Institute. These might be the key reasons for getting different miRNAs. Twenty-two of the identified miRNAs were differentially expressed in leaf rust susceptible and resistant NILs under mock and pathogen inoculated conditions. Many of the identified miRNAs were expressed more in susceptible plants in response to the rust pathogen in an effort to prevent the spread of the pathogen to adjacent tissues. The predicted miRNA targets included genes coding for structural proteins, metabolically active proteins, transcription factors, transcriptional activators, transcriptional regulators, and transporters, besides defense signaling and stress responsive proteins. Identification and functional characterization of entire miRNAs and their targets will lay the groundwork to unravel the complex miRNA-mediated regulatory networks involving the mysteries of host-pathogen interactions.

#### 5. Conclusions

We identified 52 miRNAs that are found for the first time in wheat but conserved in other plants including one novel wheat specific miRNA. The wheat specific miRNA could not be ascertained to any known miRNA family, whereas the remaining 51 miRNAs were distributed into 19 miRNA families. Twenty-two miRNAs were differentially expressed in response to leaf rust pathogenesis. Based on the computational target prediction and GO analysis of the target

genes, we found that many of these miRNAs were involved in development and biotic and abiotic stresses, thereby providing useful information regarding their regulatory roles in plant physiological processes like defense and signaling. Most putative target transcripts were involved in several biological processes related to nucleic acid binding (transcription factors) and catalytic activities. Insights into the miRNAs and their target genes will greatly contribute to the deeper understanding of posttranscriptional gene regulation in response to leaf rust induced biotic stress in wheat. The identification and characterization of miRNAs from wheat will also enrich the miRNA repertoire and provide aid in miRNA-based research on leaf rust pathogenesis in wheat.

## Abbreviation

|            |  |
|------------|--|
| AMFE:      | Adjusted minimal folding free energy                 |
| BLAST:     | Basic Local Alignment Search Tool                    |
| EST:       | Expressed sequence tags                              |
| GO:        | Gene ontology  |
| MFE:       | Minimal folding free energy                          |
| MFEI:      | Minimal folding free energy index                    |
| miRNA:     | microRNA   |
| NCBI:      | National Center for Biotechnology Information        |
| NIL:       | Near-isogenic line                                   |
| Pre-miRNA: | precursor miRNA                                      |
| SAGE:      | Serial analysis of gene expression                   |
| siRNA:     | Small interfering RNA                                |
| SOLiD:     | Sequencing by Oligonucleotide Ligation and Detection |
| TamiRNA:   | <i>Triticum aestivum</i> microRNA                    |
| WEGO:      | web gene ontology annotation                         |
| RuBisCO:   | Ribulose-1,5-bisphosphate carboxylase/oxygenase.     |

## Conflict of Interests

The authors declare that there is no conflict of interests.

## Acknowledgments

The authors are grateful to the Department of Biotechnology, Government of India (Grant no. BT/PR6037/AGR/02/308/05), for providing financial support, Centre of Excellence, Technical Education Quality Improvement Program-II (Grant no. NPIU/TEQIP II/FIN/31/158), and Department of Agriculture, Government of Jharkhand (5/B.K.V/Misc/12/2001), for infrastructure development fund. Fellowships to Dhananjay Kumar provided by the Council of Scientific and Industrial Research [9/554 (0026) 2010-EMR-I] are gratefully acknowledged.

## References

- [1] F.A.O. Food and Agriculture Organization of the United Nations, 2012, <http://www.fao.org/home/en/>.
- [2] B. S. Gill, R. Appels, A. Botha-Oberholster et al., "A workshop report on wheat genome sequencing: international genome research on wheat consortium," *Genetics*, vol. 168, no. 2, pp. 1087–1096, 2004.
- [3] R. Brenchley, M. Spannagl, M. Pfeifer et al., "Analysis of the bread wheat genome using whole-genome shotgun sequencing," *Nature*, vol. 491, pp. 705–710, 2012.
- [4] R. Dean, J. A. L. Van Kan, Z. A. Pretorius et al., "The Top 10 fungal pathogens in molecular plant pathology," *Molecular Plant Pathology*, vol. 13, no. 4, pp. 414–430, 2012.
- [5] R. A. McIntosh and Z. A. Pretorius, "Borlaug global rust Initiative provides momentum for wheat rust research," *Euphytica*, vol. 179, no. 1, pp. 1–2, 2011.
- [6] E. J. Chapman and J. C. Carrington, "Specialization and evolution of endogenous small RNA pathways," *Nature Reviews Genetics*, vol. 8, no. 11, pp. 884–896, 2007.
- [7] S. Li, L. Liu, X. Zhuang et al., "MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in *Arabidopsis*," *Cell*, vol. 153, no. 3, pp. 562–574, 2013.
- [8] F. Vazquez, "Arabidopsis endogenous small RNAs: highways and byways," *Trends in Plant Science*, vol. 11, no. 9, pp. 460–468, 2006.
- [9] A. A. Millar and P. M. Waterhouse, "Plant and animal microRNAs: similarities and differences," *Functional and Integrative Genomics*, vol. 5, no. 3, pp. 129–135, 2005.
- [10] M. Megraw and A. G. Hatzigeorgiou, "MicroRNA promoter analysis," *Methods in Molecular Biology*, vol. 592, pp. 149–161, 2010.
- [11] Y. Lee, M. Kim, J. Han et al., "MicroRNA genes are transcribed by RNA polymerase II," *The EMBO Journal*, vol. 23, no. 20, pp. 4051–4060, 2004.
- [12] D. P. Bartel, "MicroRNAs: genomics, biogenesis, mechanism, and function," *Cell*, vol. 116, no. 2, pp. 281–297, 2004.
- [13] D. Baulcombe, "RNA silencing in plants," *Nature*, vol. 431, no. 7006, pp. 356–363, 2004.
- [14] S. Campo, C. Peris-Peris, C. Siré et al., "Identification of a novel microRNA (miRNA) from rice that targets an alternatively spliced transcript of the Nramp6 (*Natural resistance-associated macrophage protein 6*) gene involved in pathogen resistance," *New Phytologist*, vol. 199, no. 1, pp. 212–227, 2013.
- [15] R. Sunkar, V. Chinnusamy, J. Zhu, and J. Zhu, "Small RNAs as big players in plant abiotic stress responses and nutrient deprivation," *Trends in Plant Science*, vol. 12, no. 7, pp. 301–309, 2007.
- [16] M. Xin, Y. Wang, Y. Yao et al., "Diverse set of microRNAs are responsive to powdery mildew infection and heat stress in wheat (*Triticum aestivum* L.)," *BMC Plant Biology*, vol. 10, article 123, 2010.
- [17] T. Unver, D. M. Namuth-Covert, and H. Budak, "Review of current methodological approaches for characterizing MicroRNAs in plants," *International Journal of Plant Genomics*, vol. 2009, Article ID 262463, 11 pages, 2009.
- [18] B. H. Zhang, X. P. Pan, Q. L. Wang, G. P. Cobb, and T. A. Anderson, "Identification and characterization of new plant microRNAs using EST analysis," *Cell Research*, vol. 15, no. 5, pp. 336–360, 2005.
- [19] E. Bonnet, J. Wuyts, P. Rouzé, and Y. Van De Peer, "Detection of 91 potential conserved plant microRNAs in *Arabidopsis thaliana* and *Oryza sativa* identifies important target genes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 31, pp. 11511–11516, 2004.
- [20] Y. S. Han, F. L. Luan, H. L. Zhu et al., "Computational identification of microRNAs and their targets in wheat (*Triticum*

- aestivum* L.),” *Science in China C*, vol. 52, no. 11, pp. 1091–1100, 2009.
- [21] B. Zhang, X. Pan, C. H. Cannon, G. P. Cobb, and T. A. Anderson, “Conservation and divergence of plant microRNA genes,” *Plant Journal*, vol. 46, no. 2, pp. 243–259, 2006.
- [22] M. Din and M. Y. K. Barozai, “Profiling and characterization of eggplant (*Solanum melongena* L.) microRNAs and their targets,” *Molecular Biology Reports*, vol. 41, no. 2, pp. 889–894, 2014.
- [23] J. R. Brown and P. Sanseau, “A computational view of microRNAs and their targets,” *Drug Discovery Today*, vol. 10, no. 8, pp. 595–601, 2005.
- [24] 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.
- [25] R. Sunkar, X. Zhou, Y. Zheng, W. Zhang, and J. Zhu, “Identification of novel and candidate miRNAs in rice by high throughput sequencing,” *BMC Plant Biology*, vol. 8, article 25, 2008.
- [26] D. Klevebring, N. R. Street, N. Fahlgren et al., “Genome-wide profiling of *Populus* small RNAs,” *BMC Genomics*, vol. 10, article 620, 2009.
- [27] V. Pantaleo, G. Szittyta, S. Moxon et al., “Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis,” *Plant Journal*, vol. 62, no. 6, pp. 960–976, 2010.
- [28] G. Szittyta, 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.
- [29] D. Qiu, X. Pan, I. W. Wilson et al., “High throughput sequencing technology reveals that the taxoid elicitor methyl jasmonate regulates microRNA expression in Chinese yew (*Taxus chinensis*),” *Gene*, vol. 436, no. 1–2, pp. 37–44, 2009.
- [30] F. Yi, S. Xie, Y. Liu, X. Qi, and J. Yu, “Genome-wide characterization of microRNA in foxtail millet (*Setaria italica*),” *BMC Plant Biology*, vol. 13, article 212, 2013.
- [31] B. Wei, T. Cai, R. Zhang et al., “Novel microRNAs uncovered by deep sequencing of small RNA transcriptomes in bread wheat (*Triticum aestivum* L.) and *Brachypodium distachyon* (L.) Beauv,” *Functional and Integrative Genomics*, vol. 9, no. 4, pp. 499–511, 2009.
- [32] B. C. Meyers, M. J. Axtell, B. Bartel et al., “Criteria for annotation of plant microRNAs,” *Plant Cell*, vol. 20, no. 12, pp. 3186–3190, 2008.
- [33] T. Tanaka, F. Kobayashi, G. P. Joshi et al., “Next-Generation survey sequencing and the molecular organization of wheat chromosome 6B,” *DNA Research*, vol. 21, no. 2, pp. 103–114, 2014.
- [34] A. Bipinraj, B. Honrao, M. Prashar, S. Bhardwaj, S. Rao, and S. Tamhankar, “Validation and identification of molecular markers linked to the leaf rust resistance gene *Lr28* in wheat,” *Journal of Applied Genetics*, vol. 52, no. 2, pp. 171–175, 2011.
- [35] D. Singh, G. Bhaganagare, R. Bandopadhyay, K. V. Prabhu, P. K. Gupta, and K. Mukhopadhyay, “Targeted spatio-temporal expression based characterization of state of infection and time-point of maximum defense in wheat NILs during leaf rust infection,” *Molecular Biology Reports*, vol. 39, no. 10, pp. 9373–9382, 2012.
- [36] G. Hu and F. H. J. Rijkenberg, “Scanning electron microscopy of early infection structure formation by *Puccinia recondita* f. sp. *tritici* on and in susceptible and resistant wheat lines,” *Mycological Research*, vol. 102, no. 4, pp. 391–399, 1998.
- [37] T. E. Coram, M. L. Settles, and X. Chen, “Transcriptome analysis of high-temperature adult-plant resistance conditioned by *Yr39* during the wheat *Puccinia striiformis* f. sp. *tritici* interaction,” *Molecular Plant Pathology*, vol. 9, no. 4, pp. 479–493, 2008.
- [38] M. S. Stark, S. Tyagi, D. J. Nancarrow et al., “Characterization of the melanoma miRNAome by deep sequencing,” *PLoS ONE*, vol. 5, no. 3, Article ID e9685, 2010.
- [39] S. Griffiths-Jones, H. K. Saini, S. Van Dongen, and A. J. Enright, “miRBase: tools for microRNA genomics,” *Nucleic Acids Research*, vol. 36, no. 1, pp. D154–D158, 2008.
- [40] A. Kozomara and S. Griffiths-Jones, “MiRBase: integrating microRNA annotation and deep-sequencing data,” *Nucleic Acids Research*, vol. 39, no. 1, pp. D152–D157, 2011.
- [41] Y. Yao, G. Guo, Z. Ni et al., “Cloning and characterization of microRNAs from wheat (*Triticum aestivum* L.),” *Genome Biology*, vol. 8, article R96, 2007.
- [42] C. Xue, F. Li, T. He, G. Liu, Y. Li, and X. Zhang, “Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine,” *BMC Bioinformatics*, vol. 6, article 310, 2005.
- [43] M. Zuker, “Mfold web server for nucleic acid folding and hybridization prediction,” *Nucleic Acids Research*, vol. 31, no. 13, pp. 3406–3415, 2003.
- [44] V. Ambros, B. Bartel, D. P. Bartel et al., “A uniform system for microRNA annotation,” *RNA*, vol. 9, no. 3, pp. 277–279, 2003.
- [45] B. H. Zhang, X. P. Pan, S. B. Cox, G. P. Cobb, and T. A. Anderson, “Evidence that miRNAs are different from other RNAs,” *Cellular and Molecular Life Sciences*, vol. 63, no. 2, pp. 246–254, 2006.
- [46] K. Y. Kurtoglu, M. Kantar, S. J. Lucas, and H. Budak, “Unique and conserved MicroRNAs in wheat chromosome 5D revealed by next-generation sequencing,” *PLoS ONE*, vol. 8, no. 7, Article ID e69801, 2013.
- [47] 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.
- [48] M. B. Stocks, S. Moxon, D. Mapleson et al., “The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets,” *Bioinformatics*, vol. 28, no. 15, pp. 2059–2061, 2012.
- [49] S. Götz, J. M. García-Gómez, J. Terol et al., “High-throughput functional annotation and data mining with the Blast2GO suite,” *Nucleic Acids Research*, vol. 36, no. 10, pp. 3420–3435, 2008.
- [50] J. Ye, L. Fang, H. Zheng et al., “WEGO: a web tool for plotting GO annotations,” *Nucleic Acids Research*, vol. 34, pp. W293–W297, 2006.
- [51] R. Margis, A. F. Fusaro, N. A. Smith et al., “The evolution and diversification of dicers in plants,” *FEBS Letters*, vol. 580, no. 10, pp. 2442–2450, 2006.
- [52] M. Wang, Q. Wang, and B. Wang, “Identification and characterization of microRNAs in asiatic cotton (*Gossypium arboreum* L.),” *PLoS ONE*, vol. 7, no. 4, Article ID e33696, 2012.
- [53] L. Yang and S. He, “A bioinformatics-based update on microRNAs and their targets in rainbow trout (*Oncorhynchus mykiss*),” *Gene*, vol. 533, no. 1, pp. 261–269, 2014.
- [54] 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.
- [55] I. L. Hofacker, M. Fekete, and P. F. Stadler, “Secondary structure prediction for aligned RNA sequences,” *Journal of Molecular Biology*, vol. 319, no. 5, pp. 1059–1066, 2002.

- [56] J. Krol, I. Loedige, and W. Filipowicz, "The widespread regulation of microRNA biogenesis, function and decay," *Nature Reviews Genetics*, vol. 11, no. 9, pp. 597–610, 2010.
- [57] M. W. Rhoades, B. J. Reinhart, L. P. Lim, C. B. Burge, B. Bartel, and D. P. Bartel, "Prediction of plant microRNA targets," *Cell*, vol. 110, no. 4, pp. 513–520, 2002.
- [58] V. Eldem, S. Okay, and T. Unver, "Plant microRNAs: new players in functional genomics," *Turkish Journal of Agriculture and Forestry*, vol. 37, pp. 1–21, 2012.
- [59] R. Sunkar, Y. F. Li, and G. Jagadeeswaran, "Functions of microRNAs in plant stress responses," *Trends in Plant Science*, vol. 17, no. 4, pp. 196–203, 2012.
- [60] 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.
- [61] 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.
- [62] T. Unver, I. Parmaksiz, and E. Dündar, "Identification of conserved micro-RNAs and their target transcripts in opium poppy (*Papaver somniferum* L.)," *Plant Cell Reports*, vol. 29, no. 7, pp. 757–769, 2010.
- [63] L. Williams, C. C. Carles, K. S. Osmont, and J. C. Fletcher, "A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the *Arabidopsis* *ARF2*, *ARF3*, and *ARF4* genes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9703–9708, 2005.
- [64] E. Allen, Z. Xie, A. M. Gustafson, G. 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.
- [65] J. T. Cuperus, N. Fahlgren, and J. C. Carrington, "Evolution and functional diversification of MIRNA genes," *Plant Cell*, vol. 23, no. 2, pp. 431–442, 2011.
- [66] L. Zhou, Y. Liu, Z. Liu, D. Kong, M. Duan, and L. Luo, "Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*," *Journal of Experimental Botany*, vol. 61, no. 15, pp. 4157–4168, 2010.
- [67] F. Meng, H. Liu, K. Wang et al., "Development-associated microRNAs in grains of wheat (*Triticum aestivum* L.)," *BMC Plant Biology*, vol. 13, p. 140, 2013.
- [68] V. Eldem, U. Çelikkol Akçay, E. Ozhuner, Y. Bakir, S. Uranbey, and T. Unver, "Genomewide identification of miRNAs responsive to drought in peach (*Prunus persica*) by high-throughput deep sequencing," *PLoS ONE*, vol. 7, no. 12, Article ID e50298, 2012.
- [69] S. Lu, Y. H. Sun, H. Amerson, and V. L. Chiang, "MicroRNAs in loblolly pine (*Pinus taeda* L.) and their association with fusiform rust gall development," *Plant Journal*, vol. 51, no. 6, pp. 1077–1098, 2007.