Physiological and Molecular Plant Pathology 87 (2014) 76–85

Contents lists available at ScienceDirect

Physiological and Molecular Plant Pathology

journal homepage: www.elsevier.com/locate/pmpp

# Next-generation sequencing and micro RNAs analysis reveal SA/MeJA1/ABA pathway genes mediated systemic acquired resistance (SAR) and its master regulation via production of *phased*, *trans-acting* siRNAs against stem rot pathogen *Macrophomina phaseolina* in a RIL population of jute (*Corchorus capsularis*)

Chinmay Biswas*, Piyali Dey, P.G. Karmakar, Subrata Satpathy

*Central Research Institute for Jute and Allied Fibres (CRIJAF), Barrackpore, Kolkata, West Bengal 700120, India*

ARTICLE INFO

*Article history:*
Accepted 14 July 2014
Available online 24 July 2014

*Keywords:*
Transcriptome analysis
Next-generation sequencing
Jute RIL
SAR
miRNAs

ABSTRACT

A RIL population of jute was developed by crossing one resistant accession CIM 036 and a susceptible variety JRC 412. Two cDNA libraries were constructed using pool of mRNA from healthy as well as infected seedlings from all the 177 RIL lines. A significant number of defense genes involved in the defense-response were identified viz. cell wall biosynthesis, reactive oxygen species (ROS), salicylic acid (SA), ethylene, jasmonic acid (JA), abscissic acid (ABA), hormone signaling, hypersensitive response (HR) and programmed cell death (PCD) pathways. Furthermore, microRNA analysis revealed that Trans-acting siRNAs (tasiRNAs) negatively regulate these target transcripts and are characterized by siRNAs spaced in 21-nucleotide (nt) "phased" intervals. We identified highly abundant 22-nt miRNA families that target conserved domains in these SA/JA/ABA precursors and trigger the production of trans-acting siRNAs. SA and JA1 transcripts were found to be cleaved by these 22-nt miRNA generating phasiRNA, suggesting silencing pathogenicity pathway of *Macrophomina phaseolina.* Gene function annotation was studied in jute-*M. phaseolina* interaction in each of the 177 lines of a RIL population. tasiRNA based SAR regulation demonstrated master regulator of a large gene family. It is the first report of studying resistance mechanism in jute against *M. phaseolina* in a RIL population through transcript and miRNA analysis.



## Introduction

*Macrophomina phaseolina* (Tassi) Goid. is a devastating necrotrophic fungal pathogen with worldwide distribution which can infect more than 500 crop and non-crop plant species [1]. Major crops such as maize, sorghum, common bean, green gram, jute, cotton, soybean, sunflower and sesame are known hosts of the pathogen [2,3]. It causes stem rot disease in jute (*Corchorus olitorius*), an important bast fiber crop. The disease is prevalent in all the jute growing areas of the world with various kinds of symptoms such as damping off, seedling blight, leaf blight, collar rot, stem rot and root rot. Average yield loss due to this disease is about 10%, but it can go up to 35–40% in case of severe infection [4]. Host plant resistance-based management of *Macrophomina* is a potential option for resource-poor jute farmers. However, in jute, no variety has been found to be completely resistant to this disease which may be attributed to the semi-saprophytic nature and wide host range of the pathogen. Only minor sources of resistance have been reported among a few genotypes evaluated in India [5,6] and Bangladesh [7]. So far no studies have identified genomic regions involved in resistance against the pathogen in jute. In other crops viz., cowpea [3], sorghum [8], soybean [9] and common bean [10] potentially useful sources of resistance against *M. phaseolina* have been reported. Transcriptome is the complete set of transcripts in a cell responsible for a particular mechanism or developmental stage. Ray et al. showed variation at transcript (cDNA) level using three 13 mer primers in beta amino butyric acid (BABA) treated detached jute leaves of a susceptible variety JRC 412. They identified 24 differentially expressed transcripts, however gene function annotation and naturally occurring defense gene was not studied [11]. The purpose of the present study was to evaluate the levels of

* Corresponding author. Crop Protection Division, Central Research Institute for Jute and Allied Fibres (CRIJAF), Barrackpore, Kolkata, West Bengal, India.
*E-mail address:* chinmaybiswas@rediffmail.com (C. Biswas).

http://dx.doi.org/10.1016/j.pmpp.2014.07.003
0885-5765/

resistance against *M. phaseolina* infection in a recombinant inbred line (RIL) population of jute and to identify the underlying mechanism of resistance through transcriptome and micro RNA analysis.

For the present study we developed a RIL population of jute by crossing CIM 036 (♀) × JRC 412(♂) and cDNA libraries for non-inoculated, inoculated healthy and inoculated infected were created from all the 177 RIL lines. We employed Sanger sequencing and 454 Newbler (Roche Diagnostics) or SeqMan™ NGen™ to analyze the RIL transcriptomes. As plant microRNAs are known to play vital role in various processes in a post-transcriptional manner by down-regulating target gene products we also made microRNA analysis by using an Illumina HiSeq instrument to identify the novel mature miRNAs and integrated PARE (parallel analysis of RNA ends) data to reveal their cleavage sites for each annotated gene. Our findings might identify various defense genes activation in jute-*M. phaseolina* interaction and reveal the miRNA triggering mechanism therein.

## Materials and methods

### *Development of RIL population*

Previous study on screening for stem rot (caused by *M. phaseolina*) resistance in jute (*Corchorus capsularis*) accessions carried out at three different locations viz., Barrackpore (22°46′N and 88°23′E), Budbud (23°23′N and 87°33′E) and Sorbhog (26°29′ N and 90°53′E) revealed field resistance in nine accessions of *C. capsularis* [6]. Among those, an improved germplasm accession, CIM 036 showed highest resistance reaction and a variety JRC 412 showed most susceptible reaction. A RIL mapping population ($F_6$ generation) for resistance to *M. phaseolina* was developed by crossing these two genotypes CIM 036 (♀) × JRC 412(♂) during 2003–2010 at Central Research Institute for Jute and Allied Fibres (CRIJAF), Barrackpore. From $F_2$ seeds, individual plants were raised and were selfed to advance these populations to $F_6$ generation following single seed descent (SSD) method. 177 lines from the RIL population were used in the present experiment.

### *Fungal culture maintenance and inoculation*

The pathogen *M. phaseolina* was isolated from infected jute (*C. capsularis*) plant (cultivar JRC 412) at the research farm of Central Research Institute for Jute and Allied Fibres, Barrackpore, Kolkata, India. The fungus culture was maintained at 25 °C on potato dextrose agar (PDA). 3 days old fungal culture was used for challenged inoculation of 177 lines of the jute RIL. One week old jute plants were sprayed with fungal suspension containing $6.2 \times 10^3$ cfu per ml and the inoculum was prepared following the procedure described by Biswas et al. [12]. Untreated healthy plants served as control. Observations were taken three days after inoculation and the tissues were used for mRNA isolation.

### *Plant samples*

Plant samples were collected from *M. phaseolina* inoculated as well as non-inoculated (healthy) plants form all the 177 RIL lines. There were three replications for inoculated as well as non-inoculated plants for each RIL line and three plants constituted one replication. Seven days old whole seedlings were taken as samples.

### *mRNA isolation*

Total mRNA was isolated using Oligotex Direct mRNA kit (Qiagen) according to the manufacturer's protocol. RNA was treated with RNase-Free DNase (Qiagen) to digest any remaining genomic DNA. RNA was quantified using a UV-spectrophotometer and its quality and integrity were examined in 1.2% agarose gel containing ethidium bromide.

### *cDNA cloning*

Two cDNA libraries were constructed, one from the infected plants and the other from the healthy check as referral genome. Twenty six pairs of micro satellite SSR primers were used for cDNA preparation, among them fifteen SSR primers were found effective (Table S5). The libraries were prepared using the cDNA Library Construction Kit (Invitrogen, CA, USA) according to the manufacturer's instructions. Briefly, the double-stranded cDNAs were fractioned and cloned in the pTriplEx2 vector of the same kit. The library was amplified in *Escherichia coli* DH-5 cells (Invitrogen), placed on LB agar and grown overnight at 37 °C. Plasmid preparations of the individual transformants were performed in 96-well plates.

### *Sanger sequencing*

To analyze the transcriptomes of RIL population we generated Sanger sequences for about ~9000 cDNA clones from a subtractive RIL library enriched with genes highly expressed in stem rot infected tissues using capillary sequencing. A total of 8101 cDNA sequences were obtained from each RIL line after filtering reads for quality. RNA was isolated using the previously described method and was reverse transcribed using cDNA Synthesis Kit (Qiagen, CA). Subtractive RIL libraries were constructed using healthy as tester and inoculated as driver following the manufacturer's instructions (Clonetec). Sequencing of the Subtractive libraries was conducted by an automated Sanger sequencing protocol.

### *Transcript assembly and contig annotation*

The 454 sequence reads were assembled into contigs using 454 Newbler (Roche Diagnostics) or SeqMan™ NGen™ v1.2 software (DNAStar, Inc), optimized for 454 next generation data. cDNA libraries were constructed using random priming which resulted in low poly A/T tail contamination and therefore no filtering was performed. Seqman removes low quality ends including homopolymers runs of poly (A/T) that have lower qualities in 454 sequencing. Contamination with mitochondrial and chloroplast genes was assessed by running a BLASTX search against *Arabidopsis thaliana* mitochondria and chloroplast proteomes. Assembly using 454 and 9000 Sanger sequences was performed, however only the 454 sequences were compared. Programs from the EMBOSS package [13], CLUSTALW [14] and DOTTER [15] were used in sequence analysis. Pair-wise sequence alignments were produced with the program EMBOSS and WATER program was used for a gap creation penalty of 30.0 and a gap extension penalty of 0.1. For sequence assemblies, 454 contigs were converted to artificial reads assigning a Phred quality score of either 20 or 40 to each base using the CONSED package [16]. Base calling of the 454-Sanger reads was done using PHRED (v. 020425.c) [17]. Hybrid assemblies of the 454-Sanger sequences were done with PHRAP (version 0.990319) [18]. Assembled contigs were mapped to reference sequences using the MUMmer package (version 3.18, [19]). Full-length contigs were identified by running a BLASTX search against the *A. thaliana* proteome and comparing the lengths of the aligned portion of each contig and the putative proteins. The annotation of contigs was performed by BLASTX [20] against the *A thaliana* proteome (e-value = $e^{-5}$) and the Gene Ontology (GO) [21] system. Transcript assembly was performed from healthy referral genome and gene

annotation was predicted as summarized in Supplementary Table S4. Comparison of GO annotation distribution between species was conducted using the GOstat program [22] set to the following parameters: GO-DB: tair; Min Sub-GO length: 3; *P*-Value Cutoff: 0.01; GO-Cluster Cutoff: −1; with no correction for multiple testing because the high dependence between GO terms will cause the test to be over conservative. To determine model species with most best hits to *Arabidopsis* transcript contigs, BLAST alignments were conducted by querying the studied contigs against the proteomes of algae, moss and higher plant species with fully sequenced genomes *(Chlamydomonas reinhardtii*, *Physcomitrella patens*, *Selaginella moellendorffii*, *Oryza sativa*, *Vitis vinifera, Populus trichocarpa*, *Carica papaya* and *A. thaliana*) and the e-values of the best hits from each species were compared.

*Identification of DSR (defense stem rot) genes in jute tissues*

DEGseq [23] was used to identify gene specific differences in transcript abundance. The DEGseq package was chosen because it integrates several statistical methods and can estimate a theoretical replicate when an experimental one is not provided [24]. The number of 454 reads per contig for each gene was compared between inoculated and healthy stem tissues in RIL lines separately. Similar analyses were performed for gene orthologs from species. Orthologs were identified using a reciprocal best hit approach. DEGseq employs a random sampling model based on the read count in inoculated and healthy stem tissue libraries and performs a hypothesis test based on that model. Two theoretical four-fold local standard deviation lines could be drawn on the expression MA-plot to estimate the noise level of genes with different intensities and to identify gene expression differences in different libraries. Genes passing the threshold are identified as exhibiting DSR. GO enrichment analyses were performed using Blast2Go software [25].

*Validation tests of DSR by real time-PCR*

Real-time quantitative RT-PCR tests were conducted to determine the extent to which the number of EST reads per gene obtained by shotgun sequencing accurately reflected transcript levels in the source tissues. RT-PCR estimates of transcript abundance were conducted on RNA from healthy and diseased tissues from RIL population. Quantitative real time PCRs (qRT-PCRs) were prepared using the SYBR Green Master Mix kit (Applied Biosystems) and run in an Applied Biosystems 7500 Fast Real-Time PCR system with default parameters. Primers were designed using Primer Express® software (Applied Biosystems) listed in Supplementary Table S6. Different gene encoding regions were used as endogenous standards to normalize template quantity. In addition, RTPCR analyses were performed to confirm the expression of DSR already identified using in silico expression analysis. For each gene, three biological replicates (three different trees) and three technical replicates were performed. Statistical analyses used Statistica 6.0 software (StatSoft Poland Inc., Tulsa, OH, USA), to estimate the significance of the differences.

*Sequencing of small RNAs*

Micro-RNA was isolated using micro-RNA isolation kit (Invitrogen) from selected 21 resistant lines and small RNA libraries were constructed and sequenced at Illumina on an Illumina HiSeq instrument. Twenty-one small RNA libraries were made from each line. Approximately 62 million small RNA sequences were obtained after removing adapters and low-quality reads, with trimmed lengths between 18 and 34 nt. Small RNAs in leaf tissues are predominantly found in two sizes: 21 nt and 22 nt.

*Small RNA informatics analysis*

A stringent filter retained all matches with scores ≤5; scoring was assigned using the CleaveLand pipeline [26,27]. The PARE data were integrated and phasing analysis was performed. As a final check of loci with phasing scores >15, scores and abundances of small RNAs from each high-scoring locus were graphed and checked visually to remove false positives such as miRNAs with numerous low-abundance peaks that could incorrectly pass our filters. We also manually removed unannotated tRNA and rRNA-like loci with high phasing scores because of their high small RNA levels.

## Results

*Sequencing summary*

In total more than 7.2 million (72, 22,567) reads were generated corresponding to 1.5 million nucleotides of cDNA sequences from the plant species *C. capsularis*. Transcript contigs were assembled from the pyrosequencing reads using Newbler software (Roche 454) and designated version 1. A second assembly was developed based on the sequences generated from both pyrosequencing and capillary sequencing reads using SeqMan™ NGen™ v1.2 software (DNAStar, Inc). The version 2 contig set included longer contigs and greater numbers of sequences were integrated into the contigs as compared to the original 454 Newbler assemblies. The combination of Sanger sequences and 454 sequences also resulted in slightly fewer but longer contigs. General information about the sequences and contigs identified are summarized in Table S1. Analysis of the RIL population transcriptomes generated over one and half million sequencing reads and yielded a total of 93,018 contigs for the two separate assemblies of the infected and healthy cDNA libraries. A small fraction of contigs matched mitochondrial (1.3%) and chloroplast (3%) genes. The tissues had best BLASTX alignments to the *A. thaliana* proteome. Transcriptome assembly, version 2 (using all of the reads combined across all tissues) led to the identification of 48,501 contigs from all *M. phaseolina* inoculated jute RIL population. From pyrosequencing alone 48,335 contigs were generated. The Sanger sequencing also led to the identification of same numbers of contigs (Table S2). GO annotations using the *A. thaliana* proteome as reference showed that the transcriptome of this species covered a wide range of biological processes (Fig. 1). The distribution of biological processes of the identified contigs from RIL population (Fig. 1) did not show any statistically significant differences. BLASTX alignments to model system proteomes showed that ~60% of the transcript contig sequences from the RIL population studied have strong similarity to predicted proteins in *A. thaliana* or *P. trichocarpa*. The remaining contigs (~30%) did not match any sequence in *A. thaliana* or *P. trichocarpa* proteomes. We observed a bias towards longer sequences in the contigs with BLASTX alignments to the model proteomes. The distribution of contig length showed that ~85% of sequences without BLASTX alignments to the proteomes of the two model species were short (<250 nt). Over 35% of contigs from jute RIL population had best alignments to *A. thaliana* whereas only ~5% of the contigs had best alignments to *Vitis vinifera* and *P. trichocarpa* (Fig. 2). Coverage of the transcriptomes from the 48,335 contigs, a total of 11,431 and 10,016 large transcript contigs (more than 800 nucleotides) were identified from inoculated and healthy plants respectively. The size of the transcript contigs assembled including large ones ranged approximately from 258 bp to 1038 bp. About 4–12% of the jute RIL contigs

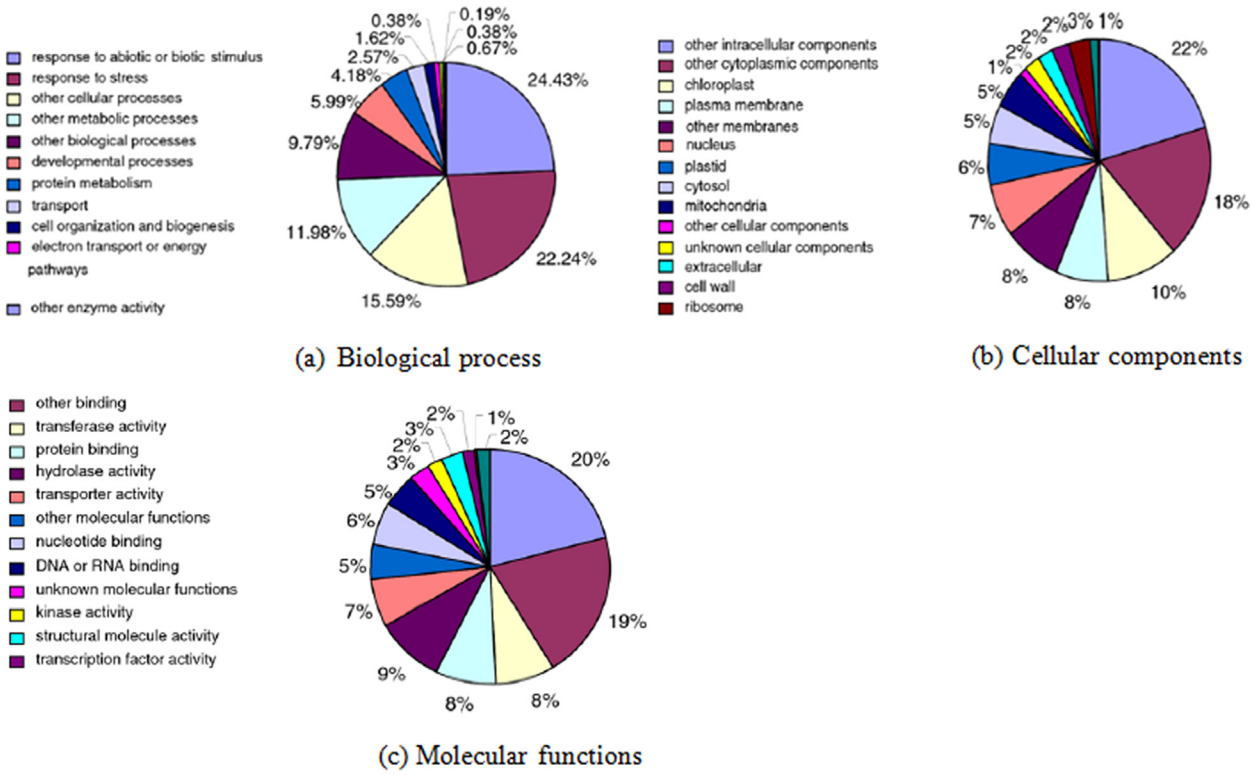


**Fig. 1.** GO analysis of differentially expressed genes in *Macrophomina phaseolina* inoculated RIL lines. The percentages of GO-terms in the categories (a) biological process (b) cellular components and (c) molecular functions are shown.

covered at least 70% of the length of the coding sequences relative to the respective genes in *A. thaliana*. Analyses of the length of contigs showed that 344 (6.0%) and 874 (6.7%) contigs were full length in RIL population. Analysis of gene family size in RIL lines as well as in *P. trichocarpa*, *A. thaliana* and *Oryza sativa* showed that the number of genes per family identified in jute is similar to their counterpart in the model plant species suggesting a good coverage of the transcriptome in *C. capsularis* species. For instance, the number of members per gene family correlates well between *C. capsularis* and the other model plant species, with correlation coefficients of $r = 0.8$ for *C. capsularis* versus *Arabidopsis* and $r = 0.74$ for *C. capsularis* versus *Populus*.

### *Defense related genes in RIL population*

In silico analysis of transcript abundance using the DEGseq approach [16] identified 1715 contigs in *M. phaseolina* inoculated asymptomatic jute plants. The number of reads per transcript contig ranged between 7 and 6388 in stem rot infected plants and between 0 and 756 in healthy check. GO annotation distribution

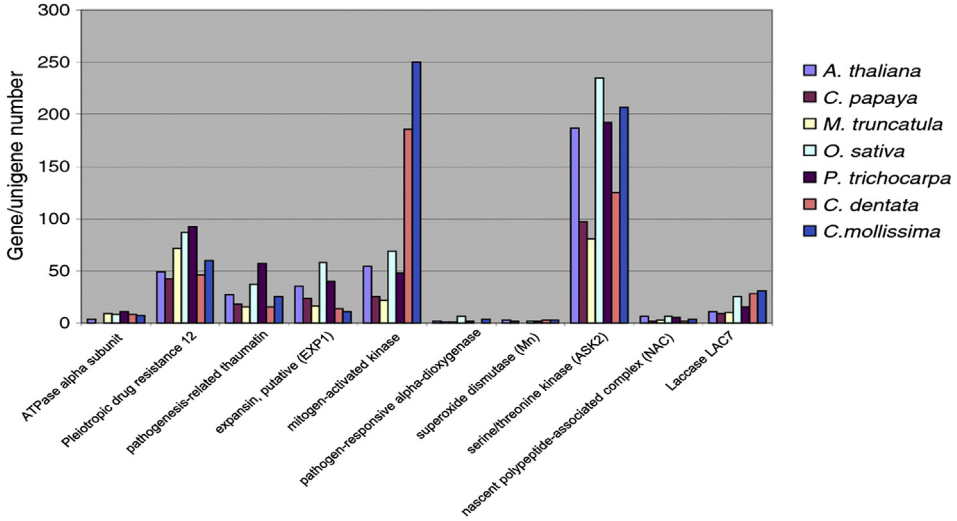


**Fig. 2.** Size of some defense related gene families in model species: *Arabidopsis thaliana*, *Populus trichocarpa*, *Oryza sativa*, *Carica papaya*, *Medicago truncatula*, *Castanea molissima* and *Castanea dentate*.

(Fig. 1) showed that 158 of the identified genes from inoculated RIL lines were involved in response to abiotic or biotic stimuli. 22 percent of genes from this functional category were involved in defense against biotic stresses. Most of the transcripts were highly abundant in stem rot tissues and thus represent good candidates for defense against the pathogen *M. phaseolina*. The most frequent molecular functions of the identified defense-related genes were hydrolases, protein binding, transferases, and transporters as revealed by GO annotation distribution (Fig. 1). Several annotation categories including "secondary metabolic process", "oxidoreductase", "cellulose and pectin containing cell wall", "hydrolases" and "lyases activity" were significantly over-represented in inoculated non-infected lines than in inoculated infected lines (Table 1). A statistical analysis using the GOstat program [28] confirmed the enrichment of inoculated non-infected lines transcriptome in these functional categories ($p$-value <0.01). On the other hand, several functional categories including mainly house-keeping genes such as "structural constituent of the ribosome", "translation", "ribosome biogenesis and assembly", and "protein metabolic process" were over-represented in healthy check than in inoculated non-infected lines ($p$-value < 0.01) (Table S4). The inoculated infected lines encoded normal cellular proteins or programmed cell death related proteins (Table S3). The over-representation of defense related genes in inoculated resistant lines could be associated with the increase in protein synthesis at the infection site for defense against the pathogen. The list of the identified defense-related genes showing defense involves several related pathways (Table S3). The first category includes genes involved in the biosynthesis of lignin and other cell wall components such as 4-coumarate CoA ligase (4CL), Cinnamyl-Alcohol Dehydrogenase (CAD), cinnamoyl CoA reductase (CCR), peroxidase, Myb transcription factor, and UDP-glucose:thiohydroximate S-glucosyltransferase. Genes involved in programmed cell death and hypersensitivity such as Myo-inositol-1-phosphate, ATPase transporter, voltage-dependent anion channel, 2-deoxy-D-arabino-heptulosonate 7-phosphate, and cysteine proteinase precursor-like protein were also identified in diseased tissues. However, one of the highly represented categories was phytohormone signaling including ethylene, jasmonic acid (JA), salicylic acid (SA), and abscisic acid (ABA). For example, transcripts of 12 genes involved in JA response were differentially abundant in resistant jute RIL lines. These include allene oxide cyclase, JAZ1, lipoxygenase, 12-oxophytodienoate reductase, 3-ketoacyl-CoA thiolase, chitinase, plastidic fatty acid desaturase, and others. Lipoxygenase, chitinase, and ACC oxidase are among genes associated with the most of the resistant lines while interaction with *Macrophomina*. Genes involved in the response to SA include alpha-dioxygenase, mitochondrial chaperoninHSP, senescence-associated gene, and others. Genes related to the ABA response include ABA 8′ hydroxylase, 26S proteasome regulatory subunit, protein phosphatase 2C, hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate (HMBPP) synthase and others. Several other genes involved were identified such as myb transcription factors, proteases and kinases.

### *Level of resistance to M. phaseolina in jute RIL lines*

Out of 177 lines of the jute RIL population screened by challenged inoculation with *M. phaseolina* hyphal suspension, 69 lines (Table 1) were found resistant. The other 108 lines were susceptible with 10–100% incidence. The lines could be grouped into three broad categories based on disease incidence. 17 lines showed low incidence (10–20%). 54 lines showed medium incidence (21–50%). The rest 37 lines were highly susceptible with 51–100% incidence.

### *Identification and validation of jute resistant lines miRNAs*

We predicted potential miRNA targets and integrated the matches with empirical cleaved mRNA data to identify valid miRNA targets. The 87 unique *C. capsularis* miRNA candidate sequences from 127 precursors that passed filters were searched against both genome and cDNA sequences. Predicted matches with penalty scores #5 (>50,000 predicted miRNA–target pairs in the genome, and >15,000 pairs from cDNAs) were combined with a PARE (parallel analysis of RNA ends) library [29] made from *C. capsularis* leaf tissue (Supplemental Table S7). The observation of several conserved miRNA–target pairs (miR154–SA, miR97–JA, miR67–ABA, and miR490–TAS3) suggested that library quality was sufficient for validation of miRNA targets; these targets exhibited precise, high abundance cleavage products at the predicted target sites. To detect novel mRNA targets in RIL lines, we used several stringent filters for the PARE data. We confirmed 144 cleavage sites from 89 genes and 30 intergenic regions, targeted by 46 different miRNAs (Supplemental Table S8). These targets include a broad set of genes not previously known to be targets of miRNAs. While most target genes were cleaved by only one miRNA at a single recognition site, we identified two target sites for miR150 in PARE library. This is significant because in the Arabidopsis TAS3 gene, two "hits" by the 21-nt miR490 are required to trigger the production of tasiRNAs [30] which triggers the SA/JA/ABA synthesis pathway.

### *Identification of tasi RNA (trans-acting siRNA)-like phasiRNA (phased siRNA) loci in resistant lines*

Eight Arabidopsis "TAS" (trans-acting siRNA) genes generate miRNA-triggered secondary siRNAs in a 21-nt "phased" pattern [31]. A refined computational approach was applied to evaluate the phasing pattern of small RNAs [32]. The small RNAs identified from this algorithm are phased but do not necessarily function in trans; therefore, it is called phased siRNA "phasiRNA". We identified tasiRNA-generating loci as TAS genes and phasiRNA-generating loci as PHAS genes. We integrated our miRNA lists, target prediction, and the PARE data to identify the triggers for the *C. capsularis* PHAS loci. We were able to identify the miRNA triggers for most PHAS loci. At least 77 of the 114 PHAS loci (68%) are triggered via single cleavage of a 22-nt miRNA; this is called "$1_{22}$" for a single-target, 22-nt miRNA trigger event. The majority of these PHAS loci were triggered by a few high-abundance 22-nt miRNAs (miR107, miR154, miR210, and miR211). There were just a few exceptions to the predominance of $1_{22}$ PHAS loci. We also identified a novel two-hit ($2_{21}$) PHAS locus (the known example i.e. TAS3) which triggers JA path way precursor- gene.

### *Twenty-two-nucleotide miRNAs as master regulators of SA/JA/ABA pathway precursor -encoding genes and generators of phasiRNAs*

A feature of the phasiRNA loci was the preponderance of SA/JA/ABA-encoding genes, in 69 resistant lines out of total 177 RIL lines. Three 22-nt miRNA families (miR154, miR210, and miR211) were found responsible for the initiation of the "phasiRNAs" which specialize in targeting SA/JA/ABA -precursor genes, with strong complementarity to the encoded kinase-2 motif, centered near a highly conserved tryptophan (W). The miR219 targets WIN1 motif of the HopW1-1-Interacting protein 1 domain. The three-member miR218 family (miR218a/b/c) target the sequences encoding well-conserved TIR1motif (responsible for NB-LLR class defense) and P-loop [33]. Two miRNA classes viz., miR218a and miR218c preferentially target myb family transcription factor genes, while miR218b almost exclusively targets SA pathway Protein kinase genes family (Fig. 3).

**Table 1**
Differentially expressed resistant genes in jute RIL lines.

| Contig | RIL line no. | Ath. BH | JIN reads | Log2 norms | Functional protein |
|---|---|---|---|---|---|
| Contig 1$^{R}$ | 78 | AT3G10920.1 | 14 | 2.32 | Manganese superoxide dismutase (MSD1) |
| Contig 2$^{R}$ | 81 | AT1G74100.1 | 99 | 3.45 | Desulfoglucosinolate sulfotransferase |
| Contig 4$^{R}$ | 82 | AT3G60160.1 | 68 | 2.33 | Member of MRP subfamily |
| Contig 6$^{R}$ | 25 | AT1G80600.1 | 29 | 6 | HopW1-1-Interacting protein 1 (WIN1) |
| Contig7$^{R}$ | 84 | AT5G60600.1 | 111 | 05 | Hydroxy-2-methyl-2-(E)- butenyl 4-diphosphate synthase |
| Contig9$^{R}$ | 11 | AT4G11260.1 | 8 | 0.71 | SCF(TIR1) mediated degradation of Aux/IAA proteins |
| Contig11$^{R}$ | 7 | AT4G37980.1 | 33 | 5.63 | Oxidoreductase activity |
| Contig12$^{R}$ | 73 | AT3G60190.1 | 294 | 5.20 | Arabidopsis dynamin-related protein 1E, DRP1E |
| Contig13$^{R}$ | 210 | AT2G26560.1 | 25 | 2.90 | Lipid acyl hydrolase with wide substrate specificity |
| Contig14$^{R}$ | 71 | AT4G37870.1 | 66 | 1.43 | Putative phosphoenolpyruvate carboxykinase |
| Contig15$^{R}$ | 1 | AT4G37870.1 | 45 | 1.78 | Putative phosphoenolpyruvate carboxykinase |
| Contig16$^{R}$ | 72 | AT4G23130.1 | 35 | 1.12 | Receptor-like protein kinase |
| Contig18$^{R}$ | 20 | AT4G09320.1 | 9 | 5.37 | Nucleoside diphosphate kinase type 1 (NDPK1) gene |
| Contig19$^{R}$ | 4 | ATCG00120.1 | 94 | 1.31 | ATPase alpha subunit |
| Contig20$^{R}$ | 17 | AT1G35720.1 | 66 | 6.44 | A member of the annexin gene family |
| Contig22$^{R}$ | 09 | AT4G11600.1 | 8 | 2.78 | Glutathione peroxidase |
| Contig25$^{R}$ | 116 | AT3G11410.1 | 23 | 0.97 | Protein phosphatase 2C |
| Contig26$^{R}$ | 12 | AT1G69530.1 | 19 | 2.40 | Member of Alpha-expansin Gene Family |
| Contig27$^{R}$ | 75 | AT4G37990.1 | 47 | 2.00 | Alcohol:NADP + oxidoreduct ase |
| Contig28$^{R}$ | 103 | AT1G04410.1 | 15 | 0.85 | Malate dehydrogenase, cytosolic, putative |
| Contig29$^{R}$ | 11 | AT4G16260.1 | 80 | 2.93 | catalytic/cation binding/hydrolase |
| Contig31$^{R}$ | 101 | AT3G57240.1 | 133 | 2.00 | Glycosyl hydrolase family 17 |
| Contig32$^{R}$ | 64 | AT3G57240.1 | 25 | 7.48 | Sigma factor |
| Contig33$^{R}$ | 85 | AT3G25070.1 | 45 | 2.33 | R protein complex |
| Contig34$^{R}$ | 15 | AT1G69530.1 | 84 | 0.85 | Member of Alpha-expansin Gene Family |
| Contig35$^{R}$ | 19 | AT5G47390.1 | 55 | 5.00 | myb family transcription factor |
| Contig36$^{R}$ | 8 | AT5G35620.1 | 19 | 3.094 | Cap-binding protein |
| Contig38$^{R}$ | 128 | AT1G02500 | 24 | 4.78 | S-adenosylmethionine synthetase |
| Contig40$^{R}$ | 36 | AT1G02800 | 171 | 3.28 | Endo-1,4-beta glucanase (CEL2) |
| Contig41$^{R}$ | 112 | AT1G05010 | 8 | 1.89 | 1-aminocyclopropane-1- carboxylate oxidase |
| Contig42$^{R}$ | 104 | AT1G13280 | 10 | 1.39 | Allene oxide cyclase |
| Contig43$^{R}$ | 14 | AT1G15520 | 28 | 3.66 | Glyceraldehyde-3-phosphate dehydrogenase C2 |
| Contig45$^{R}$ | 7 | AT1G22450 | 56 | 4.00 | Subunit 6b of cytochrome c oxidase |
| Contig46$^{R}$ | 146 | AT1G51470 | 140 | 5.25 | Myrosinase |
| Contig47$^{R}$ | 179 | AT1G58440 | 14 | 3.07 | Squalene monooxygenase activity |
| Contig48$^{R}$ | 22 | AT2G16500 | 39 | 6.05 | 12-oxophytodienoate reductase |
| Contig49$^{R}$ | 8 | AT2G18950 | 10 | 1.64 | homogentisate phytyltransferase |
| Contig50$^{R}$ | 128 | AT2G30080 | 47 | 2.90 | Member of Fe(II) transporter isolog family |
| Contig51$^{R}$ | 36 | AT2G38710 | 7 | 1.57 | AMMECR1 family |
| Contig52$^{R}$ | 112 | AT2G47510 | 12 | 1.47 | FUM1: fumarase |
| Contig53$^{R}$ | 1 | AT4G34135 | 17 | 2.40 | Flavonol 7-O-glucosyltransferase (EC 2.4.1.237) |
| Contig54$^{R}$ | 72 | AT2G33150 | 42 | 0.95 | Organellar 3-ketoacyl-CoA thiolase |
| Contig55$^{R}$ | 20 | AT3G12490 | 36 | 6.44 | Protein with an ankyrin motif |
| Contig56$^{R}$ | 4 | AT3G23600 | 39 | 2.99 | Dienelactone hydrolase family protein |
| Contig58$^{R}$ | 17 | AT3G45140 | 39 | 6.05 | Chloroplast lipoxygenase |
| Contig59$^{R}$ | 09 | AT4G01070 | 72 | 0.71 | Glycosyltransferase (UGT72B1) |
| Contig60$^{R}$ | 116 | AT4G37530 | 73 | 5.36 | Peroxidase, putative |
| Contig62$^{R}$ | 78 | AT5G06950 | 103 | 1.67 | Transcription factor of the B-ZIP family |
| Contig64$^{R}$ | 81 | AT5G14740 | 8 | 2.36 | Beta carbonic anhydrase |
| Contig65$^{R}$ | 82 | AT5G45340 | 217 | 1.72 | Protein with ABA 8′-hydroxylase activity |
| Contig66$^{R}$ | 25 | AT5G52640 | 34 | 5.37 | Cytosolic heat shock protein AtHSP90.1 |
| Contig67$^{R}$ | 84 | AT5G64250 | 11 | 1.84 | 2-nitropropane dioxygenase family/NPD |
| Contig69$^{R}$ | 82 | AT1G55020 | 7 | 1.94 | Lipoxygenase |
| Contig70$^{R}$ | 25 | AT1G45249.1 | 150 | 1.75 | Leucine zipper transcription factor |
| Contig71$^{R}$ | 84 | AT2Q40140 | 29 | 2.40 | CCCH Type zinc finger protein |
| Contig73$^{R}$ | 11 | AT75231510 | 17 | 5.20 | Halleri calmodulin-binding protein |
| Contig74$^{R}$ | 7 | AT1G02500 | 8 | 2.00 | S-adenosylmethionine synthetase |
| Contig76$^{R}$ | 73 | AT1G05010 | 147 | 0.85 | 1-aminocyclopropane-1- carboxylate oxidase |
| Contig77$^{R}$ | 210 | AT1G13280 | 184 | 4.52 | Allene oxide cyclase |
| Contig79$^{R}$ | 71 | AT1G22450 | 16 | 7.05 | Subunit 6b of cytochrome c oxidase |
| Contig80$^{R}$ | 1 | AT1G47128 | 22 | 5.20 | Cysteine proteinase precursor- like protein |
| Contig82$^{R}$ | 72 | AT1G52340 | 13 | 1.82 | Myrosinase |
| Contig83$^{R}$ | 20 | AT1G55020 | 147 | 4.15 | Lipoxygenase |
| Contig84$^{R}$ | 4 | AT2G06050 | 184 | 9.72 | 12-oxophytodienoate reductase |
| Contig85$^{R}$ | 84 | AT4G06059 | 16 | 6.20 | PR (pathogenesis-related) peptide |
| Contig87$^{R}$ | 11 | AT4Q05070 | 22 | 1.75 | Hypersensitive protein |
| Contig88$^{R}$ | 7 | AT3G10985 | 13 | 2.90 | A senescence-associated gene |
| Contig89$^{R}$ | 73 | AT3G12360 | 58 | 3.19 | Protein with an ankyrin motif |
| Contig90$^{R}$ | 210 | AT3G12490 | 17 | 7.50 | Cysteine proteinase inhibitor activity |
| Contig92$^{R}$ | 71 | AT2G18950 | 11 | 3.55 | Homogentisate phytyltransferase |
| Contig94$^{R}$ | 1 | BAB1013.1 | 49 | 5.03 | Metalo dependant hydrolase |
| Contig95$^{R}$ | 72 | AEC06958.1 | 6 | 5.00 | Phosphatase 2C and nucleotide binding kinase |
| Contig96$^{R}$ | 20 | AEC09891.1 | 15 | 1.68 | Kinase and PP2C binding domain |

(*continued on next page*)

**Table 1** (*continued*)

| Contig | RIL line no. | Ath. BH | JIN reads | Log2 norms | Functional protein |
|---|---|---|---|---|---|
| Contig97$^{R}$ | 4 | AEE 84683.1 | 59 | 2.13 | Receptor like kinase 4 |
| Contig98$^{R}$ | 17 | AEE80468.1 | 10 | 5.74 | Putative protein phosphatase 2C 51 |
| Contig99$^{R}$ | 25 | AEE78494.1 | 25 | 0.95 | U-Box domain protein containing domain family |
| Contig100$^{R}$ | 84 | AED90413.1 | 6 | 3.42 | Lucine rich repeat protein |
| Contig101$^{R}$ | 11 | AEE84800.1 | 15 | 5.74 | ATC like protein/tyrocine kinase |
| Contig103$^{R}$ | 7 | AT4Q05072 | 59 | 3.42 | Wound related protein |
| Contig104$^{R}$ | 73 | AEE7543.1 | 10 | 6.00 | SFN-1 kinase alpha subunit |
| Contig105$^{R}$ | 210 | AEE76876.1 | 25 | 4.00 | Respecter like protein kinase 2 |
| Contig106$^{R}$ | 71 | AEC06609.1 | 6 | 1.30 | Ca dependant protein kinase |
| Contig108$^{R}$ | 1 | Aee83425.1 | 6 | 0.69 | AGC −CAMP dependant protein |
| Contig119$^{R}$ | 72 | AEE33832.1 | 9 | 9.39 | S-locus lictin kinase |
| Contig111$^{R}$ | 20 | NP_194594.1 | 30 | 3.09 | Putative protein trans membrane rich protein |
| Contig112$^{R}$ | 4 | AEE86522.1 | 43 | 2.42 | Protein kinase family |
| Contig113$^{R}$ | 17 | AEE87048.1 | 63 | 6.00 | Protein kinase family |
| Contig114$^{R}$ | 09 | Aee86170.1 | 8 | 6.00 | AGC (C-AMP dependent) kinase C |
| Contig115$^{R}$ | 116 | AEE85879 | 6 | 1.04 | Protein kinase protein with adenine nucleotide alpha hydrolases-like domain |
| Contig116$^{R}$ | 12 | AEE85871.1 | 6 | 2.23 | MES 1 protein |
| Contig117$^{R}$ | 20 | AEE85870.1 | 39 | 2.70 | SA path way Protein kinase family |
| Contig118$^{R}$ | 4 | Aee84850.1 | 24 | 2.19 | Protein kinase family |
| Contig120$^{R}$ | 17 | Aee327321.1 | 8 | 3.09 | Protein kinase superfamily |
| Contig131$^{R}$ | 09 | AEE316.22 | 5 | 3.74 | Kinase domain containing protein |
| Contig17$^{R}$ | 2 | ATGO513045 | 58 | 3.19 | ITR-NBS-LRR class protein |
| Contig21$^{R}$ | 26 | N662052 | 17 | 7.70 | Putative PIR Protein |
| Contig23$^{R}$ | 79 | A1G64520 | 11 | 3.55 | Regulatory particle non ATP ase 12a (RPN12A) |
| Contig24$^{R}$ | 208 | AT1G45249.1 | 49 | 5.05 | Leucine zipper transcription factor |
| Contig30$^{R}$ | 21 | AT2G35690.1 | 6 | 5.00 | Acyl-CoA oxidase. Involved in jasmonate biosynthesis. |
| Contig44$^{R}$ | 190 | AT2G30490.1 | 15 | 1.68 | Cinnamate-4-hydroxylase |
| Contig57$^{R}$ | 109 | AT3G04720 | 59 | 2.13 | Similar to the antifungal chitin- binding |
| Contig61$^{R}$ | 102 | AT2G22240 | 10 | 5.74 | Myo-inositol-1-phosphate synthase isoform 2 |
| Contig68$^{R}$ | 106 | AT1G13440 | 25 | 0.95 | Glyceraldehyde-3-phosphate dehydrogenase C2 |
| Contig86$^{R}$ | 71 | AT1G15520 | 6 | 6.00 | ABC transporter family |
| Contig91$^{R}$ | 83 | | 6 | 6.00 | Cytosolic short-chain dehydrogenase/reductase |
| Contig102$^{R}$ | 209 | AT1G19180 | 9 | 4.00 | JAZ1 is a nuclear-localized protein |
| Contig107$^{R}$ | 16 | AEE66299.1 | 30 | 1.30 | Ca indepebdant ABA activated protein |
| Contig109$^{R}$ | 118 | AT2GO6050 | 43 | 0.69 | NADP + - isocitrate dehydrogenase |
| Contig125$^{R}$ | 76 | AT4G29040.1 | 63 | 9.39 | 26S proteasome AAA- ATPase subunit RPT2a (RPT2a) |
| Contig129$^{R}$ | 115 | AT2G16500 | 8 | 3.09 | Zinc finger protein, putative |
| Contig130$^{R}$ | 185 | AT2G06050 | 6 | 2.42 | 12-oxophytodienoate reductase |
| Contig3$^{R}$ | 30 | AT2G35690.1 | 6 | 6.00 | Acyl-CoA oxidase. Involved in jasmonate biosynthesis |
| Contig5$^{R}$ | 19 | AT3G12500.1 | 39 | 1.04 | Basic chitinase involved |
| Contig8$^{R}$ | 6 | AT1G74100.1 | 22 | 2.23 | Desulfoglucosinolate sulfotransferase |
| Contig10$^{R}$ | 201 | AT2G41430.1 | 39 | 2.70 | Hypothetical protein |
| Contig37$^{R}$ | 28 | AT2G41430 | 24 | 2.19 | Hydrophilic protein lacking Cys residues |
| Contig39$^{R}$ | 4 | AT4G37530 | 8 | 2.19 | Auxin −repressed protein like protein |
| Contig63$^{R}$ | 123 | AT4G37980 | 5 | 3.09 | Cysteine proteinase precursor- like protein |
| Contig72$^{R}$ | 137 | AT1G52340 | 6 | 3.00 | Cytosolic short-chain dehydrogenase/reductase |
| Contig75$^{R}$ | 86 | AT2G38870 | 31 | 1.08 | MADS-Box protein SOC 1 (protein suppressor of constant over expression) |
| Contig78$^{R}$ | 158 | AT5G42650 | 43 | 3.32 | Cytochrome p450 family |
| Contig81$^{R}$ | 55 | AT2G33150 | 5 | 4.74 | Salicylic acid-activated MAP kinase |
| Contig93$^{R}$ | 178 | AT3G01280 | 56 | 3.09 | Voltage-dependent anion channel |
| Contig119$^{R}$ | 96 | AT3GO4120 | 4 | 4.42 | Cytosolic GADPH (C subunit) |
| Contig121$^{R}$ | 180 | AT4G154080 | 5 | 4.74 | Sinapic acid:UDP-glucose glucosyltransferase |
| Contig122$^{R}$ | 42 | AT4G154088 | 48 | 2.12 | 2-nitropropane dioxygenase family/NPD family |
| Contig123$^{R}$ | 57 | AT4G37980 | 6 | 3.00 | ELICITOR-ACTIVATED GENE 3-1 (ELI3-1) |
| Contig124$^{R}$ | 156 | AT4G39980 | 24 | 8.00 | 2-deoxy-D-arabino-heptulosonate 7-phosphate synthase |
| Contig126$^{R}$ | 166 | AT1G02500 | 4 | 3.42 | S-adenosylmethionine synthetase |
| Contig127$^{R}$ | 191 | AT1G02800 | 20 | 1.78 | Endo-1,4-beta glucanase (CEL2) |
| Contig128$^{R}$ | 199 | AT1G20030 | 27 | 3.85 | Pathogenesis-related thaumatin family protein |
| Contig132$^{R}$ | 59 | AT4G38660 | 5 | 3.15 | Thaumatin, putative |

$^{R}$ = resistant lie, Ath.BH = *Arabidopsis* best hit, $J^{UN}$ = un inoculated healthy RIL lines, $J^{IN}$ = inoculated infected RIL lines, Log2 norm = expression ratio normalized $^{I}$ = differentially expressed inoculated infected lines $^{H}$ = differentially expressed healthy lines. JIN = inoculated non infected RIL lines.

## Discussion

### *Transcriptome analysis*

Transcriptome analysis is of growing importance in understanding biological processes in plant development as well as responses to environmental stresses [34]. Stem rot caused by *M. phaseolina* is a devastating disease in jute, but the genes responsible for its resistance are not yet known. Therefore, we took an RNA profiling approach to examine the jute - *M. phaseolina* interaction in a RIL population of jute developed by crossing one resistant accession CIM 036 and a susceptible variety JRC 412. BLASTX alignments to the proteomes of two model systems (*Arabidopsis* or *Populus*) showed that ~ 60% of the transcript contig sequences in the present study have strong similarity to predicted proteins. A large fraction of short sequences that may originate from 3′ or 5′ untranslated regions tend to be highly divergent among species. Part of these sequences may also correspond to

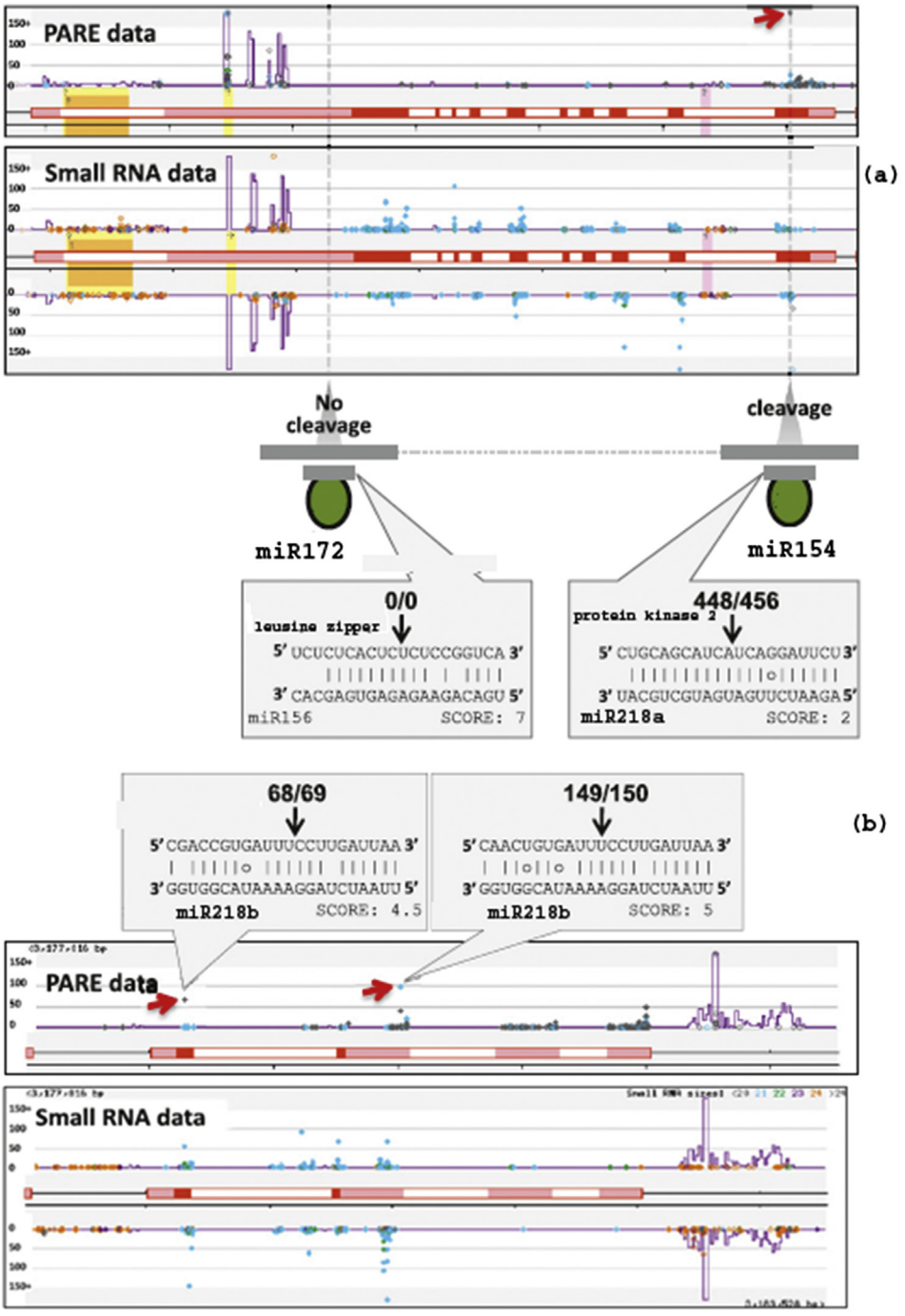


**Fig. 3.** Identification of novel class tasi RNA in jute RIL line. (a) Example of a $2_{21}$ TAS locus that encodes an SA/JA precursor path. The top panel shows the PARE data with a high abundance tag from the cleaved site (red arrowhead); for space reasons, only the coding strand data are shown for the PARE tags. The small RNA data are below; colored dots indicate small RNA sizes, with light blue indicating 21-mers. Other features are as described for the PARE images. The bottom section illustrates the predicted non cleaving miR172 site and the cleaved miR156 site, along with alignments of those miRNAs with the MeJA2 transcript and the PARE tag abundances. (b) An example of $2_{22}$ TAS locus (protein kinase 2); double cleavage by the 22-nt miRNA miR154 triggers phasiRNAs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

non-coding RNAs or potential jute-specific genes. BLAST searches against the proteomes of eight plant species with complete genome sequences (*V. vinifera*, *C. papaya*, *Medicago truncatula*, *O. sativa*, *P. trichocarpa*, *S. moellendorffii*, *P. patens*, and *C. reinhardtii)* showed that a large fraction of jute (*C. capsularis*) contigs have alignments with genes in woody species such as *V. vinifera, P. trichocarpa* and *C. papaya* versus the non-woody species. In silico analysis of gene expression identified over two fold defense genes in resistant lines upon challenged inoculation. However, most of these gene sets actually induced housekeeping genes associated with the increased resource utilization for plant defense at the infection site. The numbers of genes belonging to the category "response to biotic and abiotic stimuli" in inoculated uninfected (resistant) RIL lines are 14 and 6 times greater than in healthy control and in inoculated infected lines respectively. However, many new genes associated with jute stem rot were identified in this study. Defense related

genes identified in this study belong to well known plant pathways such as phenyl propanoid metabolism, phytohormone (JA, ABA, ethylene and SA) signaling, cell wall biosynthesis, proteolysis and others. These genes and pathways function at different times in the plant response to the pathogen. The category of genes involved in phenyl propanoid metabolism act in early plant defense serving to inhibit or to block the penetration and the progression of the pathogen. This category includes genes for biosynthesis of monolignol and other phenolic compounds. Previous studies [35,36] showed that lignin biosynthesis is crucial for cell wall deposition, one of the first lines of plant defense against the invading fungus. Besides lignin, the biosynthesis of other polymers such as callose seems to follow infection as suggested by the increased transcript abundance of UDP-glucose:thiohydroximate S-glucosyltransferase [37]. Other phenolic products that are involved in plant defense against pest and pathogens seem to be produced as well, as deduced by the presence of transcripts encoding genes such as flavanone 3-hydroxylase and flavonol 7-O-glucosyltransferase known to regulate flavonoid biosynthesis [38]. The second most important category of genes detected in response to stem rot infection includes genes from phytohormone signaling pathways including JA, SA, and ethylene. These hormones trigger the activation of induced systemic resistance and systemic acquired resistance (SAR) to necrotrophic pathogens [39,40]. The SAR is an effective defense mechanism against a broad range of pathogens and insects. Several genes from the JA response pathway such as methyl jasmonate esterase (MES1), acyl-CoA oxidase, a phyB pathway and ATPase transporter were identified [41]. Genes involved in SA response such as hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate, HopW1-1-Interacting protein 1 (WIN1) were identified [42]. The SA pathway, which is considered one of the major pathways involved in defense against necrotrophic pathogens, regulates the expression of defense effector genes and systemic acquired resistance through the repression of the auxin signaling pathway [43]. Another hormone that seems to play a role in the resistance of stem rot is abscissic acid (ABA). While ABA was described as a susceptibility factor, other studies [44] showed that it activates plant defense by priming for callose deposition or by restricting the progression of the fungus *Cochliobolus miyabeanus* in the mesophyll of rice [45]. Other signaling genes involved in SAR that induce numerous defense genes include apoplastic lipid transfer protein, basic chitinase etc [46]. The third category of genes with stem rot tissues includes genes involved in early response as part of the HR. Among these are transcripts encoding proteins such as ATPase transporter, kinases, carbonic anhydrase, AMMECR1, MIPS1, voltage-dependent anion channel, 2- deoxy-D-arabino-heptulosonate 7-phosphate (DAHP) synthase and glutathione peroxidase that were reported previously to be involved in the hypersensitivity resistance (HR) and cell death in plants under pathogenic attack [47]. Reactive oxygen species (ROS) seems to be induced following *M. phaseolina* infection as several genes involved in oxidative stress (alpha-dioxygenase, fumarase, cytosolic GADPH (C subunit), cytosolic ascorbate peroxidase APX1) had more abundant transcripts. Furthermore, several pathogenesis related (PR) genes such as elicitor-activated gene 3-1 (ELI3-1), aromatic alcohol:NADP + oxidoreductase, thaumatin, pathogenesis- related and antifungal chitin-binding protein had differentially abundant transcripts in diseased versus healthy tissues. PR proteins, of which some have antimicrobial functions [48] are mainly induced in localized pathogen attack around HR lesions. It is not yet known what roles the HR and cell death play in jute defense against a necrotrophic pathogen such as *M. phaseolina*. Alternatively, some of the genes involved in the HR may activate a systemic response of the plant or the pathogen may trigger HR to facilitate its colonization in the plant as reported for other pathogens [49]. Several other genes involved in defense such as eIF(iso)4E [50] were also implicated. The candidate genes identified in this study represent a valuable resource for studying the genetic basis underlying resistance to *Macrophomina* and the identification of the fungal pathogen resistance genes. Several candidate genes identified in this study are in the process of being analyzed for their function using transformation in planta. These resources could be used for developing SNP and SSR markers. This suggests that some of the genes identified in this study have potential to play a major role in plant defense against the stem rot disease.

### *phasiRNAs and their general roles in planta*

We identified a large number of loci that fit with the original two-hit (21-nt trigger, or "$2_{21}$") or single-hit (22-nt trigger, or "$1_{22}$") models for tasiRNA biogenesis [51]. For example, we identified a new "two-hit" plant PHAS locus, demonstrating that this pathway is not unique to TAS3, but a well-conserved unique plant developmental regulatory circuit. In fact, evidence of phased siRNAs generated from multiple target sites in a transcript have been reported for the PPR-encoding targets of TAS2 siRNAs [31,52]. The present study has expanded our understanding of tasiRNA triggering mechanism. Although tasiRNA is called so because of its activity in trans, there is evidence of cis activity in TAS3, and their functional self-targeting conserved small RNA in *Arabidopsis* [26,53]. So, the term phasiRNA is more suited to describe these secondary siRNAs as used in rice, *Arabidopsis*, etc. [54–56]. In both, the two-hit models and a single-hit model for tasiRNA biogenesis, cleavage occurs at only one site; the two "hits" of the two-hit model include one uncleaved target site [30,51]. In our study, a new class was represented by a $2_{22}$ PHAS gene. PhasiRNAs at kinase protein 2 gene are triggered after double cleavage by a 22-nt miRNA. Because 22-nt miRNAs trigger phasiRNA production in the poly-A-proximal fragment of PHAS transcripts. We observed high-abundance small RNAs between the two target sites with both, the 59 cleavage sites (downstream phased) and the 39 cleavage sites (upstream phased), which indicates a synergistic effect of two adjacent 22-nt target sites. This suggests that a $2_{22}$ cleavage product is processed inwardly from the two cleavage sites in both cap- and poly-A-proximal directions, consistent with the direction of processing for both the single-hit and two-hit models [51].

### *SA/JA/ABA defense related miRNAs have evolved in unique ways*

We identified a large number of novel miRNAs, and the analysis of these miRNAs and their targets have substantially expanded our understanding of defense small RNA biology in planta. We identified novel miRNAs and validated novel host pathogen interaction targets. By integrating PARE data with small RNA data and novel bioinformatics analyses, we identified 42 new miRNA candidates from 51 precursors in 21 resistant RIL lines. Our analysis demonstrated that *C. capsularis* genome encodes a larger set of 22-nt miRNAs. This size class of miRNAs has an innate ability to trigger phased small RNA cascades in SA/JA/ABA mediated natural SAR resistance [52,57]. The 22-nt miRNAs are produced from at least 21 loci in the *C. capsularis* genome whereas Arabidopsis generates only a few known 22-nt mature miRNAs [57], most of which are weakly expressed. Many of these jute 22-nt miRNAs are highly abundant in the tissues that we characterized. Most of the new 22-nt miRNAs we identified have large validated target sites. These 22-nt miRNAs function to trigger phasiRNAs in host tissues. New natural defense classes of miRNAs produced by *C. capsularis* and *M. phaseolina* interaction and the concerned triggering mechanism were not known previously.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.pmpp.2014.07.003.

## References


[1] Mihail JD, Taylor SJ. Interpreting variability among isolates of macrophomina phaseolina in pathogenicity, pycnidium production, and chlorate utilization. Can J Bot 1995;73:1596–603.
[2] Islam MS, Haque MS, Islam MM, Emdad EM, Halim A, Hossen QMM, et al. Tools to kill: genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina*. BMC Genomics 2012;13:493.
[3] Muchero W, Ehlers JD, Close TJ, Roberts PA. Genic SNP markers and legume synteny reveal candidate genes underlying QTL for *Macrophomina phaseolina* resistance and maturity in cowpea (*Vigna unguiculata* (L) Walp.). BMC Genomics 2011;12:8.
[4] Roy A, De RK, Ghosh SK. Diseases of bast fibre crops and their management. In: Karmakar PG, Hazra SK, Ramasubramanian T, Mandal RK, Sinha MK, Sen HS, editors. Jute and allied fibre updates. Kolkata: Central Research Institute for Jute and Allied Fibres; 2008. p. 217–41.
[5] De DK, Kaiser SAKM. Genetic analysis of resistance to stem rot pathogen (*Macrophomina phaseolina*) infecting jute. Pesq Agropec Bras Bras 1991;26(7):1017–22.
[6] Mandal RK, Sarkar S, Saha MN. Field evaluation of white jute (Corchorus capsularis L.) germplasm against *Mcrophomina phaseolina* (Tassi.) goid under Sorbhog condition. Environ Ecol 2000;18(4):814–8.
[7] Haque MMG, Mustafa MS, Islam AS. Preliminary studies of advanced generations of some intraspecific and interspecific hybrids along with their parents of corchorus for resistance to stem rot. Dacca Univ Stud Part B 1979;27:123–8.
[8] Rao DNV, Shinde VK. Inheritance of charcoal rot resistance in sorghum. J Maharashtra Agric Univ 1985;10:54–6.
[9] Smith GS, Carvil ON. Field screening of commercial and experimental soybean cultivars for their reaction to *Macrophomina phaseolina*. Plant Dis 1997;81:363–8.
[10] Songa W, Hillocks RJ, Mwango'mbe AW, Buruchara R, Konno WK. Screening common bean accessions for resistance to charcoal rot (*Macrophomina phaseolina*) in Eastern Kenya. Exp Agric 1997;33:459–68.
[11] Ray R, Ghosh A, Bera A, Dutta N, Chattopadhyaay DJ, Chakrabarti K. Analysis of differentially expressed transcripts in jute upon fungal infection and beta-amino butyric acid treatment. Physiol Mol Plant Pathol 2011;76:59–66.
[12] Biswas C, Dey P, Satpathy S, Sarkar SK, Bera A, Mahapatra BS. A simple method of DNA isolation from jute (*Corchorus olitorius*) seed suitable for PCR based detection of the pathogen *Macrophomina phaseolina* (Tassi) Goid. Lett Appl Microbiol 2013;56:105–10.
[13] The EMBOSS package. %3ca href=http://emboss.sourceforge.net/.
[14] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 1994;22:4673–80.
[15] Sonnhammer EL, Durbin R. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 1995;167(1–2). GC1-10.
[16] Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res 1998;8(3):195–202.
[17] Ewing B, Hillier LD, Wendl MC. Green P: base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 1998;8(3):175–85.
[18] PHRAP. A program for assembling shotgun DNA sequence data. http://www.phrap.org/.
[19] Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. Genome Biol 2004;5(2):R12.
[20] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215(3):403–10.
[21] Consortium GO. The gene ontology project in 2008. Nucleic Acids Res 2008;36:D440–4.
[22] Beissbarth T, Speed TP. GOstat: find statistically overrepresented gene ontologies within a group of genes. Bioinformatics 2004;20(9):1464–5.
[23] Wang L, Feng Z, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 2010;26(1):136–8.
[24] Alsford S, Turner DJ, Obado SO, Sanchez-Flores A, Glover L, Berriman M, et al. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. Genome Res 2011;21(6):915–24.
[25] Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 2008;36(10):3420–35.
[26] Allen E, Xie Z, Gustafson AM, Carrington JC. MicroRNA directed phasing during trans-acting siRNA biogenesis in plants. Cell 2005;121:207–21.
[27] Addo-Quaye C, Miller W, Axtel MJ. CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics 2009;25:130–1.
[28] Bovie C, Ongena M, Thonart P, Dommes J. Cloning and expression analysis of cDNAs corresponding to genes activated in cucumber showing systemic acquired resistance after BTH treatment. BMC Plant Biol 2004;4:15.
[29] German MA, Pillay M, Geong DH, Hetawal A, Luo S, Janardhanan P, et al. Global identification of microRNA- target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 2008;26:941–6.
[30] Axtell MJ, Jan C, Rajagopalan R, Bartel DP. A two-hit trigger for siRNA biogenesis in plants. Cell 2006;127:565–77.
[31] Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, et al. Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. Plant Cell 2007;19:926–42.
[32] De Paoli E, Dorantes-Acosta A, Zhai J, Accerbi M, Jeong DH, Park S, et al. Distinct extremely abundant siRNAs associated with cosuppression in petunia. RNA 2009;15:1965–70.
[33] Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. Plant J 1990;20:317–32.
[34] Cheong YH, Chang HS, Gupta R, Wang X, Zhu T, Luan S. Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis. Plant Physiol 2002;129:661–77.
[35] Sibout R, Eudes A, Mouille G, Pollet B, Lapierre C, Jouanin L, et al. Cinnamyl alcohol dehydrogenase-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. Plant Cell 2005;17(7):2059–76.
[36] Qi X, Bakht S, Qin B, Leggett M, Hemmings A, Mellon F, et al. A different function for a member of an ancient and highly conserved cytochrome P450 family: from essential sterols to plant defense. Proc Natl Acad Sci U S A 2006;103(49):18848–53.
[37] Clay NK, Adio AM, Denoux C, Jander G, Ausubel FM. Glucosinolate metabolites required for an arabidopsis innate immune response. Science 2009;323(5910):95–101.
[38] Nicholson RL, Hammerschmidt R. Phenolic compounds and their role in disease resistance. Annu Rev Phytopathol 1992;30:369–89.
[39] Feys BJ, Parker JE. Interplay of signaling pathways in plant disease resistance. Trends Genet 2000;16(10):449–55.
[40] Glazebrook J. Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu Rev Phytopathol 2005;43:205–27.
[41] Vlot AC, Liu PP, Cameron RK, Park SW, Yang Y, Kumar D, et al. Identification of likely orthologs of tobacco salicylic acid-binding protein 2 and their role in systemic acquired resistance in *Arabidopsis thaliana*. Plant J 2008;56(3):445–56.
[42] Lee MW, Jelenska J, Greenberg JT. Arabidopsis proteins important for modulating defense responses to Pseudomonas syringae that secrete HopW1-1. Plant J 2008;54(3):452–65.
[43] Gil MJ, Coego A, Mauch-Mani B, Jorda L, Vera P. The arabidopsis csb3 mutant reveals a regulatory link between salicylic acid-mediated disease resistance and the methyl-erythritol 4-phosphate pathway. Plant J 2005;44(1):155–66.
[44] Wiese J, Kranz T, Schubert S. Induction of pathogen resistance in barley by abiotic stress. Plant Biol 2004;6(5):529–36.
[45] De Vleesschauwer D, Yang Y, Cruz CV, Hofte M. Abscisic acid-induced resistance against the brown spot pathogen *Cochliobolus miyabeanus* in rice involves MAP kinase-mediated repression of ethylene signaling. Plant Physiol 2010;152(4):2036–52.
[46] Zander M, La Camera S, Lamotte O, Metraux JP, Gatz C. Arabidopsis thaliana class-II TGA transcription factors are essential activators of jasmonic acid/ethylene-induced defense responses. Plant J 2010;61(2):200–10.
[47] La Camera S, Balague C, Gobel C, Geoffroy P, Legrand M, Feussner I, et al. The arabidopsis patatin-like protein 2 (PLP2) plays an essential role in cell death execution and differentially affects biosynthesis of oxylipins and resistance to pathogens. Mol Plant Microbe Interact 2009;22(4):469–81.
[48] Sels J, Mathys J, De Coninck BM, Cammue BP, De Bolle MF. Plant pathogenesis-related (PR) proteins: a focus on PR peptides. Plant Physiol Biochem 2008;46(11):941–50.
[49] Govrin EM, Levine A. The hypersensitive response facilitates plant infection by the necrotrophic pathogen *Botrytis cinerea*. Curr Biol 2000;10(13):751–7.
[50] Miyoshi H, Okade H, Muto S, Suehiro N, Nakashima H, Tomoo K, et al. Turnip mosaic virus VPg interacts with arabidopsis thaliana eIF(iso)4E and inhibits in vitro translation. Biochimie 2008;90(10):1427–34.
[51] Allen E, Howell MD. miRNAs in the biogenesis of transacting siRNAs in higher plants. Semin Cell Dev Biol 2010;21:798–804.
[52] Chen HM, Chen LT, Patel K, Li YH, Baulcombe DC, Wu SH. From the cover: 22-nucleotide RNAs trigger secondary siRNA biogenesis in plants. Proc Natl Acad Sci U S A 2010;107:15269–74.
[53] Jagadeeswaran G, Zheng Y, Li YF, Shukla LI, Matts J, Hoyt P, et al. Cloning and characterization of small RNAs from medicago truncatula reveals four novel legume-specific microRNA families. New Phytol 2009;184:85–98.
[54] Johnson C, Kasprzewska A, Tennessen K, Fernandes J, Nan GL, Walbot V, et al. Clusters and superclusters of phased small RNAs in the developing inflorescence of rice. Genome Res 2009;19:1429–40.
[55] Axtell MJ. Classification and comparison of small RNAs from plants. Annu Rev Plant Biol 2013;64:137–59.
[56] Fei Q, Xia R, Meyers BC. Phased, secondary, small interfering RNAs in post-transcriptional regulatory networks. Plant Cell 2013;25:2400–15.
[57] Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, et al. Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. Nat Struct Mol Biol 2010;17:997–1003.