Coordination of MicroRNAs, PhasiRNAs, and NB-LRR Genes in Response to a Plant Pathogen: Insights from Analyses of a Set of Soybean Rps Gene Near-Isogenic Lines

Meixia Zhao, Chunmei Cai, Jixian Zhai, Feng Lin, Linghong Li, Jacob Shreve, Jyothi Thimmapuram, Teresa J. Hughes, Blake C. Meyers, and Jianxin Ma*

Abstract
Disease-related genes, particularly the nucleotide binding site (NB)—leucine-rich repeat (LRR) class of R plant genes can be triggered by microRNAs (miRNAs) to generate phased small interfering RNAs (phasiRNAs), which could reduce the transcript levels of their targets. However, how global changes in NB-LRR transcript levels coordinate with changes in miRNA and phasiRNA levels in defense responses remains largely unknown. Here, we investigated changes in the relative abundance of small RNAs (sRNAs), with a focus on miRNAs and phasiRNAs and their potential targets in response to the pathogen Phytophthora sojae in the susceptible soybean [Glycine max (L.) Merr.] ‘Williams’ and nine resistant near-isogenic lines (NILs), each carrying a unique resistance to P. sojae [Rps] gene. In total, 369 distinct miRNAs, including 78 new ones, were identified in the 10 soybean lines. The majority of miRNAs were downregulated by the pathogen. Of the 525 NB-LRR genes found in the soybean reference genome, 257 were predicted to be the targets of eight abundant miRNA families and 126 (dubbed phasi-NB-LRRs or pNLs) were predicted to have produced phasiRNAs. Upregulation of 15 phasi-NB-LRRs was associated with downregulation of their corresponding phasiRNAs in the NILs; these phasiRNAs were predicted to regulate 75 additional NB-LRRs in trans. In addition, we identified putative 24-nucleotide (nt) phasiRNAs from transposons, possibly representing a novel general epigenetic mechanism for regulation of transposon activity under biotic stresses. Together, these observations suggest that miRNAs and phasiRNAs play an important role in response to plant pathogens through complex, multiple layers of post-transcriptional regulation.

Small RNAs, including miRNAs and small interfering RNAs (siRNAs), which are typically 20 to 24 nt in length, regulate gene expression at either the transcriptional or post-transcriptional level in plants and animals (Carrington and Ambros, 2003; Bartel, 2004; Carthew and Sontheimer, 2009). In plants, miRNAs are mainly 21 or 22 nt long and are processed by DICER-LIKE 1 from single-stranded MIRNA precursors with stem-loop secondary structures. Heterochromatic siRNAs are mainly 24 nt long and originate from double-stranded RNAs that rely on the RNA-dependent RNA polymerase 2 and DICER-LIKE 3 (Xie et al., 2004). MicroRNAs repress gene expression through direct cleavage of messenger RNAs or translational repression (Khraiwesh et al., 2010). In contrast, siRNAs are involved in silencing repetitive genomic regions through de novo DNA methylation and

Abbreviations: CNI, coiled-coil NB-LRR; DEG, differentially expressed gene; DEM, differentially expressed miRNA; DEP, differentially expressed phasiRNA; DEPL, differentially expressed PHAS loci; LRR, leucine-rich repeat; miRNA, microRNA; NB, nucleotide binding site; nt, nucleotide; PARE, parallel analysis of RNA end libraries; PHAS, phasiRNA producing; phasiRNA, phased small interfering RNA; pNL, phasi-NB-LRR; pTLs PHAS locus within transposable elements; siRNA, small interfering RNA; sRNA, small RNA; tasRNA, trans-acting siRNA; TAS, tasRNA producing; TIR, Toll/interleukin-1 receptor; TNL, TIR NB-LRR.
chomatin modification (Aufsatz et al., 2002; Lippman and Martienssen, 2004; Vaucheret, 2006). Recent work has demonstrated that some miRNAs can trigger the biogenesis of secondary siRNAs, termed “trans-acting siRNAs” (tasiRNAs) or phasiRNAs through the RNA-dependent RNA polymerase 6–DICER-LIKE 4 pathway (Peragine et al., 2004; Vazquez et al., 2004; Allen et al., 2005; Yoshikawa et al., 2005; Zhai et al., 2011). Phased siRNAs arise from transcripts targeted by 21-nt miRNAs with dual cleavage sites (the “two-hit” model, Axtell et al., 2006) or transcripts targeted by 22-nt miRNAs with single cleavage sites (the “one hit” model, Chen et al., 2010; Cuperus et al., 2010). These tasiRNAs or phasiRNAs are generally clustered in a 21-nt phased register starting at the initiation cleavage sites (Chen et al., 2007; Howell et al., 2007). Similar to miRNAs, phasiRNAs function in a homology-dependent manner to suppress the expression of their target genes (Zhai et al., 2011; Li et al., 2012a; Shivaprasad et al., 2012).

In the past several years, tasiRNAs or phasiRNAs and their producing loci (TAS and PHAS, respectively) have been identified in a number of plants (Allen and Howell, 2010; Fei et al., 2013). In Arabidopsis thaliana (L.) Heynh., four TAS gene families comprising eight noncoding loci are known and well characterized (reviewed by Fei et al., 2013). Of these families, TAS1, TAS2, and TAS4 are each initiated by a 22-nt miRNA trigger associated with ARGONAUTE1 through the “one-hit” model and TAS3 is targeted by the 21-nt, ARGONAUTE7-bound miR390 through the “two-hit” model (Axtell et al., 2006; Rajagopal et al., 2006; Allen and Howell, 2010; Chen et al., 2010; Cuperus et al., 2010). In addition to noncoding TAS or PHAS loci, a significant number of protein-coding genes, including genes encoding pentatricopeptide repeat proteins (Chen et al., 2007; Howell et al., 2007), transcription factors [MYB, NAC, Toll/interleukin-1 receptor (TIR)–AFB] (Si-Amour et al., 2011; Xia et al., 2012, 2013; Zhu et al., 2012), and Ca²⁺–ATPase (Wang et al., 2011), as well as NB-LRRs (Howell et al., 2007; Zhai et al., 2011; Shivaprasad et al., 2012; Kallman et al., 2013; Xia et al., 2013), have been recently found to generate phasiRNAs.

Among the phasiRNA-producing, protein-encoding loci, NB-LRRs are of particular interest. This is because NB-LRRs are the most common plant disease-resistance genes, which recognize the activity of pathogen effectors to initiate effector-triggered immunity following the gene-for-gene interaction model. Based on the N-terminal domains, NB-LRRs in plants can be classified into TIR NB-LRR (TNL) and coiled-coil NB-LRR (CNL) subclasses (Dangl and Jones, 2001; Jones and Dangl, 2006). The combination of these two NB-LRR subclasses comprises one of the largest gene families in plants. The generation of phasiRNAs from NB-LRR genes was reported in a few flowering plants (Howell et al., 2007; Zhai et al., 2011; Li et al., 2012a; Shivaprasad et al., 2012; Xia et al., 2012; Zhang et al., 2012). Because phasiRNAs target NB-LRR transcripts in cis or in trans at other NB-LRR loci, the regulatory effects of miRNA triggers on NB-LRR expression could be reinforced by the resulting phasiRNAs (Zhai et al., 2011; Li et al., 2012a; Shivaprasad et al., 2012; Xia et al., 2013).

Soybean is one of the most economically important leguminous crops with capacity to fix atmospheric N by intimate symbioses with soil bacteria known as rhizobia. A recent study showed that an R gene (a TNL) was involved in the control of host specificity in the soybean–rhizobia symbiosis and nodulation (Yang et al., 2010). Another functionally validated R gene in soybean is Rps1-k, which encodes a CNL protein that confers resistance to Phytophthora sojae, an oomycete pathogen that causes root and stem rot in soybean (Gao et al., 2005; Gao and Bhat-tacharyya, 2008). These observations suggest that similar recognition mechanisms are likely to be involved in symbiotic and pathogenic host–microbe interactions. As such, understanding the R gene-mediated molecular responses to biotic stresses in soybean is of great importance. It has been demonstrated that overexpression or suppression of a few specific miRNAs can result in decreased or increased levels of the corresponding NB-LRR transcripts (Li et al., 2012a; Shivaprasad et al., 2012). However, it remains unknown which NB-LRR genes are regulated in a complex plant genome and how global changes in NB-LRR levels coordinate with changes in miRNA and phasiRNA levels in their host plants in responses to biotic stresses.

Previously, we analyzed the transcriptomes of nine resistant soybean NILs, each carrying a unique Rps gene or allele (Rps1-a, Rps1-c, Rps1-k, Rps3-a, Rps3-b, Rps3-c, Rps4, Rps5, or Rps6) in the Williams genetic background, along with the susceptible Williams that was used as the recurrent parent to develop these NILs, before and after inoculation with the P. sojae race 1 (avirulent toward NILs, virulent toward Williams) (Lin et al., 2014). We characterized genes and multiple branches of the putative regulatory networks associated with Rps-mediated resistance to P. sojae (Lin et al., 2014). In that study, we also observed substantial variation of differentially expressed genes (DEGs), including those involved in ethylene, jasmonic acid, reactive oxygen species, and mitogen-activated protein-kinase signaling among these soybean NILs, and proposed that different timing and robustness in defense signaling may be largely responsible for such variations (Lin et al., 2014).

In the present study, we examined sRNA regulation associated with soybean resistant and susceptible responses to P. sojae in soybean, with a focus on global changes in levels of the miRNAs; transcripts of the phasiRNA-producing loci, particularly NB-LRR genes; and the resulting phasiRNAs, as well as their potential interplays in response to this particular pathogen. Both sRNA and mRNA data were generated from the nine NILs and Williams. Through characterization and comparison of the sRNAs, and mRNA data from these NILs before and after inoculation with P. sojae, we demonstrate coordinated regulatory roles of miRNAs and phasiRNAs, as well as NB-LRRs that are apparently nonresistant to the pathogen P. sojae in the defense responses. We identified putative 24-nt phasiRNAs from transposable elements; the changes in the abundance of these phasiRNAs on inoculation with
the pathogen may explain the stress-induced transcriptional and transpositional activities of transposons.

MATERIALS AND METHODS

Plant Material and Generation of sRNA Data

The susceptible cultivar Williams and its nine NILs, each containing a unique Rps gene or allele, and the protocols used for inoculation of these soybean plants with P. sojae and mock-inoculation without P. sojae were as previously described by Lin et al. (2014). Total RNAs with efficient enrichment of sRNAs were isolated using the mirNeasy Mini Kit following the manufacturer’s protocol (Qiagen, Valencia, CA). The sRNA libraries were constructed using TruSeq Small RNA Sample Preparation Kit (Qiagen) and sequenced using Illumina HiScanSQ (San Diego, CA) at Purdue University Genomics Corer Facility.

Profiling of Gene Expression with RNA-Seq Data

The RNA sequencing data from the same set of soybean samples were generated previously by Lin et al. (2014) but the data were reanalyzed with the newly released annotation of the soybean reference genome (G. max version 1.1; www.SoyBase.org, accessed 15 Dec. 2014) for profiling gene expression and identification of DEGs, following the same protocol as previously described (Lin et al., 2014).

Identification and Analyses of Putative miRNAs and Their Targets

The sRNA raw sequencing data was trimmed and filtered for low-quality reads, adaptor sequences, reads shorter than 17 nt, and reads matching the P. sojae genome sequence (Tyler et al., 2006) or structural noncoding RNAs (transfer RNA, ribosomal RNA, small nuclear RNA and small nucleolar RNA) collected from Rfam (http://rfam.sanger.ac.uk, accessed 15 Dec. 2014) and GtRNAdb (Chan and Lowe, 2008). The reads passing through the filtrations were mapped to the soybean reference genome sequence (G. max version 1.1) using Bowtie (Langmead et al., 2009). Only reads matching perfectly to the reference genome sequence were kept for further analyses. The sRNA sequencing data have been deposited at the National Center for Biotechnology Information Gene Expression Omnibus under the accession number GSE56859.

All known miRNA sequences and their precursors were downloaded from miRBase version 19 (Kozomara and Griffiths-Jones, 2011). Novel miRNAs were identified following the protocols described earlier (Jeong et al., 2011; Zhai et al., 2011) with slight modifications as illustrated in Supplemental Fig. S1. Potential targets of miRNAs were predicted using a combination of two different algorithms, psRNA_Target (Dai and Zhao 2011) and TargetFinder 1.6 (http://carringtonlab.org/resources/targetfinder, accessed 15 Dec. 2014). Additionally, three publicly available parallel analysis of RNA end (PARE) libraries (Song et al., 2011; Shamimuzzaman and Vodkin, 2012; Hu et al., 2013) were used to validate the predicted miRNA targets using CleaveLand version 3 with the parameters of penalty score $\leq 7$ and $P \leq 0.05$ (Addo-Quaye et al., 2009). CentroidFold with the CONTRAfold engine (Sato et al., 2009) was used to predict the secondary structures of MIRNA precursors. Target site protein sequence logos were generated by WebLogo version 3 (http://weblogo.threeplusone.com/create.cgi, accessed 15 Dec. 2014). Heatmaps were created using the “heatmap” package integrated in R. Differentially expressed miRNA (DEMs) were determined using the following criteria: (i) The abundance of an miRNA was greater than 10 transcripts per million reads in at least one of the 20 soybean libraries and (ii) the log$_2$ ratio of the fold change of an miRNA’s abundance in the pathogen-inoculated sample to that in the mock-inoculated sample of a same soybean line was greater than 1 or less than $-1$.

RESULTS

Small RNAs from 10 Soybean Lines Before and After Inoculation with P. sojae

To understand the regulatory roles of sRNAs in Rps-mediated defense responses to the pathogen, we sequenced sRNA libraries from the same set of soybean samples as was used for transcriptomic analysis (Lin et al., 2014). After sequence trimming and filtration (see Methods), a total of 255 million reads longer than 16 nt were obtained, representing ~19 million distinct sequences (Supplemental Table S1). Of these reads, 203 million (79.7%) were perfectly mapped to the soybean reference genome sequence (Schmutz et al., 2010, G. max version 1.1), including 76 million reads matching
structural noncoding RNAs, such as ribosomal RNA, transfer RNA, small nuclear RNA, and small nucleolar RNA. These noncoding RNA reads and the other 52 million reads, including the ones imperfectly matching the soybean genome sequence and the ones matching the P. sojae genome sequence (Tyler et al., 2006), were excluded from further analysis (see Methods). The overall patterns of sRNA distribution among different libraries were similar (Supplemental Fig. S2). In each of the 20 libraries, the 21-nt and 24-nt reads were found to be the most abundant. According to what has been observed in other plants (Law and Jacobsen, 2010), it was expected that 21 nt would be the main length of miRNAs, whereas the 24 nt would be the main length of siRNAs that primarily mediate silencing of repetitive sequences. The sRNA distribution patterns with respect to their sizes and relative abundances varied among different resistant NILs and Williams (Supplemental Fig. S2).

Identification and Characterization of Known miRNA Variants and Novel miRNAs

To profile the genome-wide patterns of miRNA expression in responses to the pathogen, we identified the miRNAs present in the 20 sRNA libraries using a bioinformatics pipeline and criteria previously described (Meyers et al., 2008; Jeong et al., 2011; Zhai et al., 2011; Supplemental Fig. S1), combined with a comparison to previously annotated miRNAs deposited in miRBase (v19.0; Kozomara and Griffiths-Jones, 2011) and a set of recently identified miRNAs from soybean (Wong et al., 2014). A total of 369 putative miRNAs from 523 MIRNA genes were expressed in at least one of the 20 samples. Of these, 291 distinct mature miRNAs corresponding to 415 MIRNA genes from G. max were included in miRBase. The most abundant sRNAs produced from at least 44 of the 415 MIRNA loci identified in our study did not correspond to the mature miRNAs represented in miRBase. The most common difference between the most abundant sRNAs and the miRBase-annotated mature miRNAs was a displacement of a nucleotide in either the 5’ or 3’ end. For instance, miR2118a and -b in miRBase were annotated as 21-nt miRNAs with a summed abundance of 178 transcripts per million reads, whereas these two miRNAs were observed as 22-nt variants with an abundance of 33,058 transcripts per million reads in our dataset. This family has also been confirmed to have a predominant length of 22 nt in a few other plants recently investigated (Zhai et al., 2011; Shivaprasad et al., 2012). Similarly, miRNAs from the other 43 MIRNA loci included in miRBase were different from those defined in our study (Supplemental Table S2). These differences may reflect the sequence variations among different soybean varieties investigated, may be the outcome of distinct processing of miRNA stem-loop precursors by DICER-LIKE proteins, or may simply be caused by previous mis-annotation. Of the 369 putative miRNAs, 76, corresponding to 108 putative miRNA precursors, were newly identified in this study, including 63 novel miRNAs that could not be classified into any known miRNA families, six novel miRNAs from known miRNA precursors, and seven new miRNAs that were classified into known miRNA families (Supplemental Fig. S1 and S3 and Supplemental Table S3). Of the 63 putative novel miRNAs, 10 were further supported by analysis of their putative cleaved targets using sequencing data from three publicly available PARE libraries of soybean (Supplemental Table S4; Song et al., 2011; Shamimuzzaman and Vodkin, 2012; Hu et al., 2013), following the criteria employed by Wong et al. (2014).

Differential Expression of miRNAs among Different Soybean Lines in Responses to P. sojae

To identify miRNAs associated with molecular responses to P. sojae, the relative abundances of the 369 distinct miRNAs in mock-inoculated and pathogen-inoculated NILs and Williams were compared. The changes in the levels of expression of most miRNAs were relatively low or moderate and varied among different lines (Fig. 1A and Supplemental Fig. S4). It was notable that the overall levels of downregulation of miRNA expression in each of the nine resistant NILs (particularly the Rps3-a, Rps3-b, Rps3-c, Rps4, and Rps5 NILs) were higher than those detected in Williams (Fig. 1B and Supplemental Table S5, Student’s paired t-test, P < 0.05).

We next identified DEMs among different soybean lines on inoculation with the pathogen and compared DEMs among or between the resistant NILs and the susceptible Williams. Out of the 369 putative nonredundant miRNAs, 119 DEMs, including two (1.7%) upregulated and 117 (98.3%) downregulated ones, were detected in at least 1 of the 10 soybean lines (Table 1, Fig. 2A and Fig. 2B). The numbers of DEMs ranged from 24 to 104 among individual lines in response to the pathogen (Table 1, Fig. 2A and Fig. 2B). Of the 119 nonredundant DEMs, 89 (75%) were specifically present in at least one of the nine NILs but were absent in Williams, which were referred to as incompatible DEMs; 21 (19%) were specifically present in Williams but absent in any of the nine NILs, which were referred to as compatible DEMs. Only eight DEMs (6%), including seven downregulated ones and one upregulated one (miR5786) were shared by all the 10 soybean lines, which were referred to as shared DEMs (Fig. 2A and Fig. 2B).

Previous hierarchical cluster and heatmap analyses of DEGs with the RNA-Seq data in the resistant NILs on inoculation with the pathogen grouped them into three clusters (Lin et al., 2014; Fig. 2C). In the context of these three clusters of the NILs, the detected incompatible DEMs were analyzed by broad-range comparison, where all incompatible DEMs within a cluster were counted, and narrow-range comparison, where only the incompatible DEMs shared by all the NILs within a cluster were counted. Broad-range comparison showed that 20 incompatible DEMs were shared by the three clusters, whereas 0, 1, and 36 in compatible DEMs were unique to Cluster I (C-I), Cluster II (C-II), and Cluster III (C-III), respectively. In contrast, no incompatible DEMs shared by all the three clusters were revealed by narrow-range comparison (Fig. 2D).
Characterization of the Target Genes of miRNAs

To understand the potential regulatory roles of the detected DEMs, we predicted the target genes of all miRNAs identified in this study. These target genes were first annotated by bioinformatic approaches then further validated by using sequencing data from three publicly available PARE libraries of soybean (Song et al., 2011; Shamimuzzaman and Vodkin, 2012; Hu et al., 2013). A total of 320 genes fell into a standard set of criteria for target cleavage (penalty score \( \leq 7 \), \( P \leq 0.05 \)) using the software package CleaveLand version 3 (Supplemental Table S4; Addo-Quaye et al., 2009). Additionally, 63 targets were annotated based on their anticorrelated transcript changes in abundance with the corresponding miRNA changes (Supplemental Table S6). Of the 383 predicted target genes, 80 appeared to be upregulated by the pathogen and showed anticorrelated transcript changes in abundance with the corresponding DEMs in the majority of the 10 soybean lines (Supplemental Table S6). These include 20 NB-LRR genes; 15 transcription factor genes, such as MYB, NAC, zinc finger, bZIP, and bHLH (Supplemental Fig. S5A); 10 genes encoding oxidative stress-associated proteins such as lipoxygenase, polyphenol oxidase, copper amine oxidase, and cytochrome c oxidase; and nine protein kinase and four heat shock genes (Supplemental Fig. S5B and Supplemental Table S6). Many of these genes were proposed to be involved in defense responses to plant pathogens (Lin et al., 2014).

The differentially expressed NB-LRR genes drew our specific attention, although none of these NB-LRR genes appeared to be the Rps genes in the NILs that were directly responsible for resistance to \( P. \) sojae. In general, most of these genes were upregulated, particularly in the NILs (Supplemental Fig. S5C and Supplemental Table S6), but the patterns of their expressional changes varied among the 10 lines. For instance, Glyma15g37255.1 targeted by DEM miR1510, was differentially upregulated in all nine NILs in response to the pathogen but did not change at the transcriptional level in Williams. Glyma03g14888.1, targeted by DEM miR2109, was upregulated in six NILs (Rps3-a, Rps3-b, Rps3-c, Rps4, Rps5, and Rps6) but downregulated in Williams. Glyma02g02777.1 was differentially upregulated in all 10 soybean lines, including Williams.

Table 1. Differentially expressed microRNAs (miRNAs) in the 10 soybean lines.

| Soybean lines | Differentially expressed miRNAs (|log2FC| ≥ 1) | Upregulated miRNAs | Downregulated miRNAs |
|---------------|---------------------------------|--------------------|----------------------|
| Williams (rps) | 30                              | 2                  | 28                   |
| Rps1-a        | 60                              | 2                  | 58                   |
| Rps1-c        | 24                              | 2                  | 22                   |
| Rps1-k        | 54                              | 2                  | 52                   |
| Rps3-a        | 78                              | 1                  | 77                   |
| Rps3-b        | 89                              | 1                  | 88                   |
| Rps3-c        | 103                             | 1                  | 102                  |
| Rps4          | 96                              | 1                  | 95                   |
| Rps5          | 104                             | 1                  | 103                  |
| Rps6          | 57                              | 2                  | 55                   |

Figure 1. Changes in relative abundance of micro RNAs (miRNAs) in the 10 soybean lines before and after inoculation with Phytophthora sojae. (A) Relative abundance (scaled as log2 fold changes of transcripts per million reads (TPM)) of individual miRNA families in Williams (left), and the near-isogenic lines (NILs) with Rps3-c (right) before and after inoculation. (B) Comparison of the relative abundance of total miRNAs in the 10 soybean lines before and after inoculation.
Characterization of Putative NB-LRR PHAS Loci and phasiRNA and miRNA Triggers of NB-LRR PHAS Loci

Knowing that NB-LRR genes can be regulated by miRNA via production of phasiRNAs in plants (Zhai et al., 2011; Li et al., 2012a; Shivaprasad et al., 2012; Fei et al., 2013), we wondered whether and at what scale phasiRNAs may be produced from the NB-LRR loci. We first analyzed the 525 NB-LRR genes including 290 CNLs and 235 TNLs annotated in the soybean genome (Supplemental Table S7) and their putative miRNA regulators. Our data suggest that 349 (66.5%) NB-LRR genes were targeted by miRNAs (Fig. 3A), 257 (73.6%) of which were regulated by eight large miRNA families: miR1510, miR1507, miR2109 and miR482/2118, miR5668, miR5376, miR172, and miR5041 (Fig. 3B; Supplemental Fig. S6). The families miR1510, miR1507, miR2109, and miR482/2118 were previously predicted to target NB-LRR genes in Medicago truncatula Gaertn. and a few other nonleguminous plant species (Zhai et al., 2011; Shivaprasad et al., 2012) but such interactions have not yet been detected in soybean. The families miR5668, miR5376, miR172, and miR5041 had not been predicted to target NB-LRR genes in any species before this study. Following the methods previously described (Zhai et al., 2011), we predicted that miR1510 and miR482/2118 would both target the P-binding loop domain of 122 NB-LRRs (31 CNLs and 91 TNLs), miR172 targets the kinase 2 domain of 122 NB-LRRs (31 CNLs and 13 TNLs), MicroR5376 mainly targets 18 CNLs at the coiled coil domain and miR5668 mainly targets an unknown domain in the conserved region of 36 TNLs (Fig. 3C and Supplemental Fig. S6). The other three miRNA families (miR1507, miR5041, and miR2109) were each predicted to target either CNLs or TNLs but not both (Fig. 3C and Supplemental Fig. S6). For example, miR1507 targets the kinase 2 motif of 52 CNL genes, miR5041 target the coiled coil domain of 19 CNL genes and miR2109 targets the TIR-1 domain of 83 TNL genes. The predicted target domain of miR482/2118 was validated in tomato (Solanum lycopersicum L.) (Shivaprasad et al., 2012) and M. trunculata (Zhai et al., 2011). Among these eight large miRNA families predicted to target NB-LRR genes, miR1507, miR5041, and miR2109 were each predicted to target either CNLs or TNLs but not both (Fig. 3C and Supplemental Fig. S6).

Subsequently, a genome-wide identification of PHAS loci using the sRNA data from the 20 libraries was performed following a protocol described previously (Zhai et al., 2011; Song et al., 2012; Xia et al., 2013), with minor modifications (see Methods). A total of 208 nonredundant putative PHAS loci, including 190 annotated protein-coding genes and 18 intergenic loci, were identified (Fig. 4A; Supplemental Table S8 and S9). Of these 190 protein-coding loci, 126 (66.3%) were NB-LRR genes (termed pNLs;
Zhao et al. (2011), 23 (12.1%) were pentatricopeptide repeat superfamily genes; six (3.2%) were TIR/AFB F-box family genes, and four (2.1%) were NAC family genes. Other putative PHAS loci were low-copy or single-copy genes, such as two putative DCL2 genes and one SUPPRESSOR OF GENE SILENCING3 gene, which were believed to be involved in sRNA biogenesis (Zhai et al., 2011; Fig. 4A and Supplemental Table S9). Intriguingly, out of the 18 intergenic PHAS loci, 12 (66.7%) were found to be transposable elements, including nine mutators and three long terminal repeat retrotransposons (Supplemental Table S10). These transposon-derived PHAS loci are referred to as pTLs.

Of the 208 putative PHAS loci, 130 (62.5%), including 72 pNLs, were predicted to have the corresponding miRNA triggers detected in the 20 sRNA libraries. Based on the prediction, the majority (40.8%) of the PHAS loci were triggered by two 22-nt miRNA families miR1507 or miR482/2118 following the one-hit “122” mode (Fig. 4B and Supplemental Table S9; Zhai et al., 2011). Two PHAS loci Glyma09g03731 and Glyma15g14675 were predicted to be triggered by the 21-nt miR390 following the two-hit “221” mode (Supplemental Table S9). An intergenic PHAS locus was predicted to be targeted by the 22-nt miR1509 at two sites 892 bp apart, following a two-hit “221” mode (Supplemental Table S9). Of the 72 pNLs, 23, 23, and 19 were predicted to be triggered by miR1507, miR482/2118, and miR1510, respectively. Because, in most cases, multiple PHAS loci shared the same miRNA trigger, how the abundance of transcripts from individual PHAS loci was affected by their miRNA triggers could not be evaluated precisely.

Differential Expressions of phasiRNAs versus PHAS Loci among Different Soybean Lines in Response to *P. sojae*

To understand potential regulatory roles of phasiRNAs in response to the pathogen, we first examined the expression changes of the identified phasiRNAs before and after inoculation. Out of the 208 putative PHAS loci, 166 were predicted to have produced unique phasiRNAs.
in all of the 20 samples (Table 2). The unique 21-nt phased siRNAs from a single PHAS locus were combined to represent the relative abundance of all phased siRNAs from this locus in a particular sample and was treated as a unit to identify DEPs on inoculation with the pathogen. DEPs were found at a higher proportion of the 166 PHAS loci in most of the nine resistant NILs than in the susceptible line Williams. For example, DEPs were detected from 147 out of the 166 PHAS loci in the resistant line carrying Rps3-c in response to the pathogen, whereas DEPs were detected at only 31 PHAS loci in Williams (Table 2). Of the 164 nonredundant DEPs, 15 showed upregulation and 159 exhibited downregulation in at least 1 of the 10 soybean lines. The number of DEPs ranged from 31 to 147 among individual lines (Table 2).

Of the 164 nonredundant DEPs, 133 (81.1%) were specifically expressed in at least one of the nine NILs but not in Williams, which were referred to as incompatible DEPs; 25 (15.2%) were specifically expressed in Williams but not in any of the nine NILs, which were referred to as compatible DEPs. Only two DEPs (1.2%), both downregulated, were shared by all 10 soybean lines; these were referred to as shared DEPs (Fig. 5A). Additionally, four (2.4%) DEPs were specifically up- or downregulated in the NILs, but down- or upregulated in Williams. In the context of the three clusters of the NILs, the detected incompatible DEPs were analyzed by both broad-range comparison and narrow-range comparison as described above. Broad-range comparison showed that 33 incompatible DEPs were shared by the three clusters, 16 of which were detected from NB-LRRs, whereas 0, 3, and 41 in compatible DEPs were unique to C-I, C-II and C-III, respectively. In contrast, no incompatible DEPs shared by all three clusters were revealed by narrow-range comparison (Supplemental Fig. S7A).

In total, 87 differentially expressed PHAS loci (DEPLs), including 44 upregulated and 57 downregulated ones, were detected in at least 1 of the 10 soybean lines. The numbers of DEPLs ranged from 20 to 41 among individual lines (Fig. 5B). Of the 87 DEPLs, 52 (59.8%) were specifically expressed in at least one of the nine NILs but not in Williams, which were referred to as incompatible DEPLs; 19 (21.8%) were specifically expressed in Williams but not in any of the nine NILs, which were referred to as compatible DEPLs. Only five DEPLs (5.7%), including no downregulated ones and five upregulated one, were shared by all 10 soybean lines, referred to as shared DEPLs. Additionally, 11 (12.6%) DEPLs were specifically up- or downregulated in the NILs but down- or upregulated in

### Table 2. Phased small interfering RNA (phasiRNA) precursor (PHAS) loci producing differentially expressed 21-nucleotide phasiRNAs in the 10 soybean lines before and after inoculation with Phytophthora sojae.

<table>
<thead>
<tr>
<th>Soybean Lines</th>
<th>PHAS loci producing upregulated phasiRNAs</th>
<th>PHAS loci producing downregulated phasiRNAs</th>
<th>Total PHAS loci producing differently expressed phasiRNAs (log₂FC &gt; 1)²</th>
<th>Proportions (%) of PHAS loci producing differently expressed phasiRNAs³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Williams</td>
<td>6</td>
<td>25</td>
<td>31</td>
<td>18.67</td>
</tr>
<tr>
<td>Rps1-a</td>
<td>2</td>
<td>69</td>
<td>71</td>
<td>42.77</td>
</tr>
<tr>
<td>Rps1-c</td>
<td>4</td>
<td>34</td>
<td>38</td>
<td>22.89</td>
</tr>
<tr>
<td>Rps1-k</td>
<td>1</td>
<td>80</td>
<td>81</td>
<td>48.80</td>
</tr>
<tr>
<td>Rps3-a</td>
<td>2</td>
<td>92</td>
<td>94</td>
<td>56.63</td>
</tr>
<tr>
<td>Rps3-b</td>
<td>2</td>
<td>121</td>
<td>123</td>
<td>74.10</td>
</tr>
<tr>
<td>Rps3-c</td>
<td>0</td>
<td>147</td>
<td>147</td>
<td>88.55</td>
</tr>
<tr>
<td>Rps4</td>
<td>2</td>
<td>122</td>
<td>124</td>
<td>74.70</td>
</tr>
<tr>
<td>Rps5</td>
<td>1</td>
<td>138</td>
<td>139</td>
<td>83.73</td>
</tr>
<tr>
<td>Rps6</td>
<td>2</td>
<td>26</td>
<td>28</td>
<td>16.87</td>
</tr>
</tbody>
</table>

³Only unique 21-nt reads perfectly matching each PHAS locus were used to evaluate their relative abundance before and after inoculation. In total, 166 out of the 208 PHAS loci have unique 21-nucleotide reads in all 20 soybean libraries and are thus used for comparative analysis.

---

**Figure 4. Putative phased small interfering RNA precursor (PHAS) loci and predicted microRNA (miRNA) triggers in the 10 soybean lines.** (A) Classification of putative PHAS loci based on sequence annotation. The table lists double- or single-copy PHAS loci. (B) Numbers of phased NB-LRRs (pNLs) with known and unknown miRNA triggers.
Williams. In the context of the three clusters of the NILs, the detected incompatible DEPs were analyzed using both broad-range comparison and narrow-range comparison. Broad-range comparison showed that six iDEPLs were shared by the three clusters, whereas 2, 6, and 34 incompatible DEPLs were unique to C-I, C-II, and C-III, respectively. In contrast, only two incompatible DEPLs shared by all the three clusters were revealed by narrow-range comparison (Supplemental Fig. S7B).

To address how the expressions of the PHAS loci, particularly the pNLs, and other NB-LRR genes were regulated by the resulting phasiRNAs in cis or in trans in response to the pathogen, we further analyzed the differentially expressed pNLs, including 26 upregulated and 27 downregulated pNLs (Fig. 6, Supplemental Fig. S8 and S9). Of the 26 upregulated pNLs, 17 were incompatible DEPLs, eight were shared DEPLs, and one was a compatible DEPL. Out of these 17 upregulated incompatible DEPLs, 15 showed anticorrelated expression patterns with their corresponding DEPs. It was suggested that 75 upregulated “non-pNL” NB-LRR genes could be targeted by unique DEPs generated from these 15 incompatible DEPLs in trans. In contrast, none of the 27 downregulated incompatible DEPLs were predicted to have produced upregulated DEPs upon inoculation with the pathogen.

DISCUSSION

Downregulation of a Large Proportion of sRNAs in Defense Responses

Although upregulation of the individual sRNAs that are associated with pathogen-associated molecular pattern-triggered immunity and effector-triggered immunity against plant pathogens has been reported (Katiyar-Agarwal et al., 2006; Navarro et al., 2006; Li et al., 2010; Wong et al., 2014), the relative abundances of large proportions of sRNAs in resistant NILs and susceptible Williams, including miRNAs, the 21-nt phasiRNAs, and the 24-nt siRNAs, were reduced on inoculation with the pathogen. A similar pattern of genome-wide changes in small RNA expression was also observed in soybean infected by soybean cyst nematode (Li et al., 2012b) and in loblolly pine (Pinus taeda L.) responses to fusiform rust disease in the galled stem (Lu et al., 2007). Together, these observations suggest that downregulation of sRNA biogenesis may be a general process during pathogen or nematode attacks. This process appears to be able to relieve the sRNA-mediated silencing pathways (Yi and Richards 2007; Shivaprasad et al., 2012), resulting in activation of a subset of genes involved in plant defense responses to biotic stresses (Lin et al., 2014).

Several studies have demonstrated that plant pathogens can encode a multitude of effectors or suppressors
to subvert host immunity by inhibiting the biogenesis of sRNAs (Voinnet et al., 1999; Navarro et al., 2008; Qiao et al., 2013). Two such suppressors, *Phytophthora* suppressors of RNA silencing 1 and 2 (PSR1 and PSR2), have been identified from the oomycete plant pathogen *P. sojae* (Qiao et al., 2013). It is suggested that PSR1 may target multiple DCLs or common DCL cofactors required for biogenesis of both miRNAs and siRNAs to promote infection, whereas PSR2 probably functions at the late stage of *P. sojae* infection through suppression of the tasiRNA pathway (Qiao et al., 2013). On the other hand, the temporal suppression of miRNAs and siRNAs induced by plant pathogen effectors may also enhance *R* gene activities and, to some extent, provide a balance between resistance and susceptibility during host–pathogen interactions (Li et al., 2012a). Nevertheless, downregulation of large proportions of sRNAs were consistently observed in the resistant NILs and susceptible Williams in response to the same *P. sojae* isolate, suggesting that suppression of sRNA biogenesis may be a shared mechanism behind the common phenomenon observed in both incompatible and compatible soybean–*P. sojae* interactions.

**Specificities in Defense Responses among Individual NILs with Different Rps Genes or Alleles**

In spite of such a general pattern of changes in sRNA abundance on inoculation with *P. sojae*, the defense or immune responses to the pathogen, as reflected by the numbers and specificities of the differential expressions of miRNAs, PHAS loci and phasiRNAs (DEMs, DEPLs, and DEPs) vary dramatically among NILs with different *Rps* genes or alleles (Fig. 2 and Fig. 5; Table 1 and Table 2). These results echo our previous study, which showed a clear distinction in global transcriptomic changes (DEGs) responsive to the same pathogen within the same set of soybean NILs (Lin et al., 2014). If one believes that the detected variations of DEMs, DEPLs, DEPs, and DEGs are mainly caused, directly or indirectly, by different *Rps* genes or alleles in the NILs that share the same genetic background, and that all these *Rps* genes/alleles recognize the same or different avirulence effector gene(s) in the same *P. sojae* isolate used for inoculation in this study, then the detected distinction in the molecular immune responses mediated by different *Rps* genes or alleles would be considered to be truly substantial.

Differentially expressed genes in soybean that are responsive to *P. sojae* were found to peak at 24 h after inoculation or infection (Moy et al., 2004) and thus the soybean lines in this study were sampled at 24 h after inoculation for RNA and sRNA profiling. Therefore, dramatic and rapid changes in gene expression following inoculation with or infection by a pathogen have been observed (McDonald and Cahill, 1999; Moy et al., 2004). Recent profiling of global transcriptomic changes has revealed that only a small proportion (~5%) of DEGs were shared by soybean lines resistant to bacterial leaf pustule at 6 and 12 h after inoculation with the pathogen *Xanthomonas axonopodis* (Kim et al., 2011). A more recent study identified 54 novel miRNAs in *P. sojae*-infected or mock-treated soybean root tissues at 8 h after the treatments (Wong et al., 2014); intriguingly, only two of them were present in the 20 libraries investigated in our study (Supplemental Table S3). Because profiling of gene and sRNA expression mediated by different *R* genes in response to the same pathogen has not been conducted at multiple time points after inoculation in the same tissue of any plant, it is still unknown whether the detected variations DEMs, DEPLs, DEPs, and DEGs are primarily reflective of differential timing and robustness in immune responses and signaling mediated by different *Rps* genes or alleles. Alternatively, such variations could be partially explained by potential “dilution” effects of detectable DEMs, DEPLs, DEPs, and DEGs in the proportion of neighboring cells in the collected tissues that were not responsive to the pathogen. It is also possible that such variations can partly be attributed to the specificities of the immune responses mediated by individual *Rps* genes or alleles.

It should be noted that the relative abundance of *NB-LRR* transcripts and sRNAs in this study were solely evaluated by RNA-Seq and sRNA-Seq without validation by other experimental approaches such as Northern blot analysis and quantitative polymerase chain reaction. Whether such variations were also partially affected by the methods used in this study for profiling RNA abundance remains unclear. Nevertheless, because the majority of the miRNAs predicted to target *NB-LRRs* share high levels of sequence similarity or overlap with one another (e.g., mi1510, miR482, and miR2118) and with some of the phasiRNAs generated from the *NB-LRRs* targeted by these miRNAs, the relative abundance of these individual miRNAs and phasiRNAs may not be accurately evaluated by Northern blot analysis. In addition, the transcripts of many *NB-LRRs* including the miRNA targets and *pTNLs* may not be specifically quantified by quantitative polymerase chain reaction either, because of the high levels of sequence similarity among these genes and thus the potential nonspecific amplification of such gene sequences. These limitations leave sRNA-Seq and RNA-Seq, as perhaps, the most feasible approaches to profiling changes in the relative abundance of transcripts of *NB-LRR* genes and sRNA with “unique reads” that differ by as few as a single nucleotide. We would also like to note that, because a single miRNA, as predicted, may target more than 100 *NB-LRRs* and because these targets could be further regulated, both in cis and in trans, by phasiRNAs generated from one or multiple *pTNLs*, it remains challenging to accurately elucidate the regulatory roles of individual miRNAs and phasiRNAs in response to inoculation with the pathogen.
Abundant PHAS Loci in Soybean NILs Carrying Distinct Rps Genes or Alleles

Regardless of the nature of the detected distinction in defense responses to the pathogen among different soybean NILs, our study annotated the largest number (208) of putative PHAS loci, including 126 pNLs, in any plant that has been investigated to date. A previous study identified 41 PHAS loci, including 21 pNLs in the soybean genome (Zhai et al., 2011), 39 (95%) of which were included in our dataset (Supplemental Table S9). Given that over 13 times greater depth of sRNA reads was generated in our study than in the previous study (Zhai et al., 2011), the relative abundance of PHAS loci annotated in our study, particularly pNLs, would be partially explained by the relative abundance of sRNA reads generated from the 20 soybean tissue samples. In addition, the special set of soybean NILs and inoculation with the pathogen may also partially account for the over-representation of pNLs described in our study. Nevertheless, the majority of the PHAS loci identified in this study were also detected in small RNA libraries derived from various tissues from the vegetative and reproductive parts of soybean, including flowers, leaves, and developing nodules (Arkit et al., 2014), indicating the importance of these loci in various biological processes.

PhasiRNAs Derived from Transposons May Be Important Players in the Regulation of Transposon Activity under Stress

It is worthwhile to note that pTLs have not previously been found in any eukaryotes and thus the identification of PHAS loci in transposable elements within the soybean genome under stress from the pathogen P. sojae is an intriguing observation. It has been documented that the transcriptional and transpositional activities of transposable elements in plants can be initiated by biotic and abiotic stresses (Wessler, 1996; Grandbastien, 1998), and such activity changes can generally be explained by 24-nt siRNA-mediated epigenetic regulation (Aufsatz et al., 2002; Lippman and Martienssen, 2004; Vaucheret, 2006). Theoretically, similar to phasiRNAs from pNLs that regulate NB-LRRs both in cis and in trans, phasiRNAs from pTLs should be able to target transposons both in cis and in trans. If this is the case, then the discovery of transposon-derived 21-nt phasiRNAs and their changes in abundance may represent a new regulatory mechanism for transposon activity under environmental stresses. However, because transposable elements are primarily targeted by siRNAs, the potential roles of phasiRNAs in regulating transposon activity may not be easily revealed.

Are “Nonresistant” R Genes Involved in the Defense Responses?

The differential expression of a large number of pNLs and other NB-LRR genes responsive to P. sojae appears to be a striking observation. Although classic disease-resistance R genes, primarily NB-LRRs, play essential roles in recognizing corresponding pathogen effectors during active immune responses (Dangl and Jones, 2001; Jones and Dangl, 2006), none of the NB-LRR genes annotated in the soybean reference genome could be the Rps genes that were carried by the resistant NILs investigated in this study. It is worthwhile to point out that the soybean reference genome sequence was generated from the cultivar Williams 82, one of the nine NILs, which is believed to carry Rps1-k, the only Rps gene that has been cloned to date (Gao et al., 2005; Gao and Bhattacharyya, 2008). However, it is unexpected that the Rps1-k sequences isolated from a different source of Williams 82 (Gao et al., 2005; Gao and Bhattacharyya, 2008) were not found in the reference genome sequence. Moreover, profiling of the gene expression and identification of PHAS loci in the NILs were performed by mapping the mRNA and sRNA sequencing reads from the NILs to the reference genome sequence, which does not contain any of the Rps genes present in these NILs. Therefore, how expression of the Rps genes in these NILs may change, whether these Rps genes are targeted by miRNA triggers to produce phasiRNAs, and whether these Rps genes are regulated by miRNAs and phasiRNAs in cis or in trans in immune responses to the pathogen could not be investigated.

Are non-Rps NB-LRR genes involved in immune responses? A few lines of observation garnered from this study appear to be in favor of an affirmative answer. These observations include the relatively higher abundance of downregulated pNLs and the relatively higher proportion of pNLs generating DEPs in response to inoculation with the pathogen in individual resistant NILs compared with Williams (Table 2), the subset of pNLs showing anticorrelated expression patterns with their corresponding DEPs in NILs but not in Williams (Fig. 6), and the relatively larger number of upregulated non-pNL NB-LRRs responsive to the pathogen in the majority of individual NILs compared with Williams. As 75 NB-LRRs were predicted to be regulated in trans by phasiRNAs produced by 15 pNLs (Fig. 6), these non-Rps pNLs and other NB-LRRs may also be important players in defense and immune responses by regulating the expression of the Rps genes in the soybean NILs in trans, through sRNA-mediated regulatory pathways. Whether and how such amplification effects on defense and immune responses take place remain open questions.

CONCLUSIONS

This study revealed the coordinated regulation of miRNAs, phasiRNAs, and a surprisingly large set of NB-LRRs that are apparently nonresistant to P. sojae, in response to the pathogen, as well as the amplified effects of miRNA-mediated gene regulation by producing phasiRNAs. These observations suggest that both miRNAs and phasiRNAs play an important role in plant defense through complex and multiple layers of post-transcriptional regulation of protein-encoding genes, particularly NB-LRRs. It is also suggested that miRNA-triggered biogenesis of phasiRNAs from transposable elements may
represent a novel and perhaps general mechanism for regulation of transposon activity in plants under various biotic and/or abiotic stresses.

Supplemental Information Available
Supplemental information is included with this article.

Acknowledgments
We thank Phillip SanMiguel for assistance in sRNA sequencing and Doug Yatcilla for help with software installation and testing. This work was mainly supported by soybean check-off funds from the Indiana Soybean Alliance (grant number 205267) and partially supported by funds from the National Natural Science Foundation of China (grant number 31371647) and the Taishan Scholarship and High Level Talents Foundation of Qingdao Agricultural University (grant number 631304).

References


