Title

Identification and functional characterization of soybean root hair microRNAs expressed in response to Bradyrhizobium japonicum infection.

Journal

Plant Biotechnology Journal, 14(1)

Authors

Yan, Zhe

Hossain, Md

Valdés-López, Oswaldo

et al.

Publication Date

2016

DOI

10.1111/pbi.12387

Peer reviewed

Plant Biotechnology Journal

Plant Biotechnology Journal (2016) **14**, pp. 332–341 doi: 10.1111/pbi.12387

Identification and functional characterization of soybean root hair microRNAs expressed in response to *Bradyrhizobium japonicum* infection

Zhe Yan[1,†,‡], Md Shakhawat Hossain[1,†], Oswaldo Valdés-López[1,§], Nhung T. Hoang[1], Jixian Zhai[2], Jun Wang[1,3], Marc Libault[1,¶], Laurent Brechenmacher[1], Seth Findley[1], Trupti Joshi[4], Lijuan Qiu[3], D. Janine Sherrier[2], Tieming Ji[5], Blake C. Meyers[2], Dong Xu[4] and Gary Stacey[1,*]

[1]*Divisions of Plant Science and Biochemistry, National Center for Soybean Biotechnology, Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA*

[2]*Department of Plant and Soil Sciences and Delaware Biotechnology Institute, University of Delaware, Newark, DE, USA*

[3]*National Key Facility for Crop Gene Resources and Genetic Improvement (NFCRI), Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China*

[4]*Department of Computer Science, Informatics Institute and Christopher S. Bond Life Sciences Center, University of Missouri-Columbia, Columbia, MO, USA*

[5]*Department of Statistics, University of Missouri, Columbia, MO, USA*

Received 17 October 2014;
revised 17 March 2015;
accepted 20 March 2015.
**Correspondence* (Tel +01 5738844752;
fax 573-884-9395;
email staceyg@missouri.edu)
[†]These two authors contributed equally to the work.
[‡]Present address: Center for Carbohydrate Recognition and Signaling, Department of Molecular Biology and Genetics, Aarhus University, Gustav Wieds Vej 10, Aarhus C 8000, Denmark.
[§]Present address: Laboratorio de Bioquímica, Facultad de Estudios Superiores Iztacala, UNAM, Los Reyes Iztacala, Estado de México 54090, Mexico.
[¶]Present address: Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019, USA.

Keywords: soybean, root hair, nodulation, miRNA, gma-miR4416, gma-miR2606b.

Summary

Three soybean [*Glycine max* (L) Merr.] small RNA libraries were generated and sequenced using the Illumina platform to examine the role of miRNAs during soybean nodulation. The small RNA libraries were derived from root hairs inoculated with *Bradyrhizobium japonicum* (In_RH) or mock-inoculated with water (Un_RH), as well as from the comparable inoculated stripped root samples (i.e. inoculated roots with the root hairs removed). Sequencing of these libraries identified a total of 114 miRNAs, including 22 novel miRNAs. A comparison of miRNA abundance among the 114 miRNAs identified 66 miRNAs that were differentially expressed between root hairs and stripped roots, and 48 miRNAs that were differentially regulated in infected root hairs in response to *B. japonicum* when compared to uninfected root hairs ($P \leq 0.05$). A parallel analysis of RNA ends (PARE) library was constructed and sequenced to reveal a total of 405 soybean miRNA targets, with most predicted to encode transcription factors or proteins involved in protein modification, protein degradation and hormone pathways. The roles of gma-miR4416 and gma-miR2606b during nodulation were further analysed. Ectopic expression of these two miRNAs in soybean roots resulted in significant changes in nodule numbers. miRNA target information suggested that gma-miR2606b regulates a *Mannosyl-oligosaccharide 1, 2-alpha-mannosidase* gene, while gma-miR4416 regulates the expression of a *rhizobium-induced peroxidase 1* (*RIP1*)-like peroxidase gene, *GmRIP1*, during nodulation.

Introduction

Non-protein-coding RNAs (ncRNAs) are a class of riboregulators that play a pivotal role in the control of gene transcription or translation (Eddy, 2001). For example, microRNAs (miRNAs) regulate transcription through transcript cleavage or translational inhibition. Additionally, some ncRNAs also contribute to the spread of gene silencing through the production of secondary small interfering RNAs (Allen *et al.*, 2005; Chen, 2012). miRNAs have been intensively studied in both animals and plants. Since the first miRNA was discovered in *Caenorhabditis elegans* (Lee *et al.*, 1993), the sequences of 30 000 miRNAs from 200 species have been deposited in miRBase (http://www.mirbase.org/).

In the last few years, several published reports demonstrated specific biological roles for miRNAs in the regulation of plant development, hormone signalling, pattern formation, metabolism, plant–microbe interaction and adaptation to abiotic stress (reviewed by Chen, 2009; Ruiz-Ferrer and Voinnet, 2009). Although a number of reports have focused on the role of plant miRNAs in plant–microbe interactions, relatively few studies specifically addressed the role of miRNAs in the legume nodulation process.

Soybean [*Glycine max* (L) Merr] is a legume crop that is a major source of vegetable oil and feed for animal production. A key feature of soybean is the ability to establish a nitrogen-fixing symbiosis with the soil bacterium *Bradyrhizobium japonicum*, thereby alleviating the need for nitrogen fertilizer application. Under conditions of nitrogen deficiency, *B. japonicum* infects soybean roots through root hair cells, induces the formation of symbiotic root nodules and establishes a persistent, intracellular infection. Within the nodule, the bacteria convert atmospheric nitrogen into ammonia, which is provided to the plant to support growth (Jones *et al.*, 2007; Oldroyd, 2013; Oldroyd and Downie, 2004).

Nodule cell development in soybean is synchronous. Therefore, a number of studies identified miRNAs with potential roles during

specific phases of nodule development. In one study focused on an early stage of the symbiotic interaction, 55 miRNAs families, including 35 novel miRNAs, were identified from *B. japonicium*-inoculated soybean roots 3 h postinoculation or from mock-inoculated roots (Subramanian *et al.*, 2008). Expression levels of these miRNAs were quantified, and a subset of miRNAs was found to be differentially regulated during the early phases of nodulation (Subramanian *et al.*, 2008). More recently, with the aid of the complete soybean genome sequence, reanalysis of these small RNA data resulted in the identification of 120 novel miRNAs (Turner *et al.*, 2012). In a separate study of small RNAs from a variety of soybean tissues, including nitrogen-fixing nodules collected 7, 14 and 21 days after *B. japonicium* inoculation, 129 miRNAs, including 87 novel miRNA, were identified (Joshi *et al.*, 2010). Finally, 32 miRNAs were identified from a small RNA library constructed from more mature, 28-day-old soybean nodules (Wang *et al.*, 2009). Collectively, these experiments identified a number of miRNAs that may be involved in the regulation of nodule development.

A role for miRNAs in regulating gene expression during a variety of plant–microbe interactions is well documented (reviewed by Hewezi *et al.*, 2012; Katiyar-Agarwal and Jin, 2010; Khraiwesh *et al.*, 2012; Staiger *et al.*, 2013; Zhu *et al.*, 2013). However, relatively few studies have gone further to define specific functions for miRNAs during nodule development. Combier *et al.* (2006) reported that miR169a mediates the cleavage of *MtHAP2-1,* a transcription factor, which regulates nodule development by controlling meristem persistence in *Medicago truncatula* (*M. truncatula*). Subsequently, Boualem *et al.* (2008) demonstrated that miR166 regulated expression of an *HD-ZIPIII* transcription factor known to be essential for both root and nodule development. In a more recent report, Li *et al.* (2010) demonstrated that ectopic expression of mir482 and miR1515 resulted in an increase in nodule numbers on soybean roots, suggesting a positive role for these miRNAs in the nodulation process. Turner *et al.* (2013) reported that miR160 regulates a set of auxin response factors and inhibits nodule primordia development. We reported that gma-miR172 regulates AP2 transcription factors, which directly or indirectly inhibit the level of nodule haemoglobin, which correlated with nodule development (Yan *et al.*, 2013). Recently, Wang *et al.* (2014) expanded on our earlier study by showing that miR172 positively regulates soybean nodulation through targeting *Nodule Number Control1*, an AP2 transcription factor, which directly represses expression of the early nodulin gene, *ENOD40*.

As *B. japonicum* infection of soybean root hairs is among the earliest events in the nodulation process, we focused on the activities and functions of miRNA in root hairs after inoculation. Using high-throughput sequencing technology, we sequenced three small RNA libraries from root hairs either mock-inoculated with water (Un_RH) or inoculated with *B. japonicum* (In_RH), as well as from the corresponding *B. japonicum*-inoculated stripped roots (i.e. roots with root hairs removed). At the same time, we also generated and sequenced a parallel analysis of RNA ends (PARE) library from soybean root hairs. A total of 114 miRNAs with 405 gene targets were identified, including 22 novel miRNAs. Expression analysis identified 66 miRNAs differentially expressed between root hairs and stripped roots and 48 miRNAs that were differentially regulated after *B. japonicum* inoculation in root hairs when compared to uninfected root hairs. Among these miRNAs, we focused on functional characterization of two miRNAs, gma-miR2606b and gma-miR4416, in the soybean nodulation process. The data indicate that gma-miR2606b and gma-miR4416 positively and negatively regulate soybean nodulation, respectively, and illustrate the important role of miRNAs during the early stages of soybean nodulation.

Results

Sequencing of small RNAs in soybean root hairs infected by *B. japonicum*

In our previous research, such as transcriptome, proteome, phosphoproteome and metabolome studies (Brechenmacher *et al.*, 2012, 2010; Libault *et al.*, 2010b; Nguyen *et al.*, 2012), we examined the plant root hair response to *B. japonicum* inoculation from 0 to 48 h. These experiments identified a large number of genes, proteins and metabolites regulated upon *B. japonicum* infection. Microscopic studies also indicate that the first 48 h define the period of *B. japonicum* infection of root hairs. Thus, to examine the activities of miRNAs during the early stages of *B. japonicum* infection, small RNA libraries were generated from root hairs collected and pooled at 12, 18, 24 and 48-h after mock inoculation (Un_RH) or inoculation with *B. japonicum* (In_RH). The comparable *B. japonicum*-inoculated stripped root samples (i.e. inoculated roots with the root hairs removed) were also collected. These three small RNA libraries were then sequenced using Illumina sequencing technology, resulting in the generation of 3 697 659; 1 900 537; and 3 828 884 primary sequence tags from the Un_RH, In_RH and stripped roots libraries, respectively. Mapping of these primary sequence tags to the soybean genome resulted in 637 069 (17%), 418 868 (22%) and 1 804 706 (47%) tags matching perfectly from the Un_RH, In_RH and stripped root libraries, respectively (Table 1).

The distribution of primary sequence tags of different lengths is shown for each library in Figure 1, revealing two major peaks at 21 and 24 nt. The proportion of 24-nt small RNAs was similar for the Un_RH and In_RH libraries, with 26% and 33%, respectively. The distribution in the stripped root library was different with the 24-nucleotide category, representing 56% of the total, a much larger proportion than Un_RH and In_RH. There were a higher proportion of 21-nt small RNAs in the root hair libraries with 23% and 19% in the Un_RH and In_RH libraries, respectively, and only 13% in the stripped root library. These findings are consistent with previous reports that different tissues can show different relative levels of small RNA sizes (Jeong *et al.*, 2011; Kasschau *et al.*, 2007).

The first nucleotide of the 5′ end of the small RNA sequence is characteristic of the interaction with a specific AGRONAUTE

Table 1 Number of sequenced tags matching the soybean genome, tRNA, rRNA and distinct tags in three small RNA libraries (Un_RH, In_RH and stripped root)

	Un_RH	In_RH	Stripped root
Total sequences	3 697 659	1 900 537	3 828 884
Genome-matched reads	637 069	418 868	1 804 706
Distinct genome-matched reads	178 666	150 193	800 136
Reads matched tRNA, rRNA or snoRNA	1 238 993	638 398	758 231

Un_RH and In_RH, uninfected and infected root hairs.

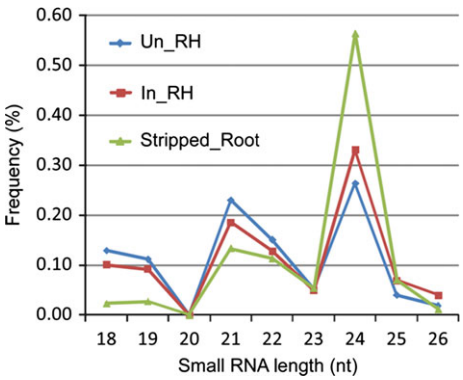

Figure 1 Size distribution of small RNAs identified from mock-inoculated root hairs (Un_RH) and *Bradyrhizobium japonicum*-infected root hairs (IN_RH) and stripped roots. Reads were normalized to transcripts per million. The size of small RNAs was plotted against frequency in each library. The majority of the tags were 21, 22 and 24 nucleotides in length.

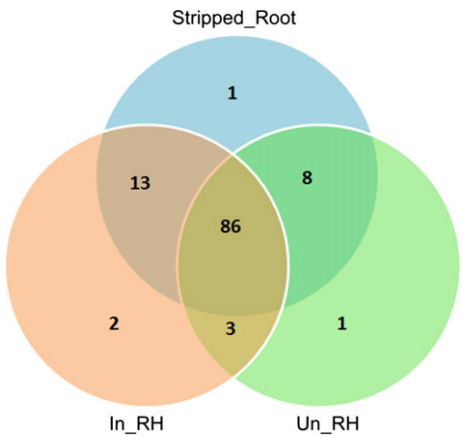

Figure 2 Venn diagram of number of miRNAs expressed in stripped roots, Un_RH and In_RH. Normalized expression levels of miRNAs from stripped roots, Un_RH and In_RH libraries were compared. Among 86 miRNAs from three libraries, one, two and one miRNAs indentified only in stripped roots, In_RH and Un_RH libraries, respectively.

(AGO) protein (Mi *et al.*, 2008). To determine whether a bias exists in the 5′ nucleotide of the various soybean small RNA classes, the distribution of the first nucleotide of each small RNA sequence was calculated relative to the length of the small RNA (Table S1). Uridine, which is characteristic of small RNA that associate with AGO1, was found to be more abundant than the other three possible nucleotides at the first nucleotide in 21- and 22-nt small RNAs, with an average of 34% and 36%, respectively. This was true of all three libraries. In contrast, 38% of the 24-nt small RNAs had adenosine as the 5′ nucleotide, which is a characteristic of small RNA that associates with AGO4. Chi-squared statistical analysis revealed a significant difference between the distribution of the 5′ nucleotide with the null hypothesis that A (adenine), G (guanine), U (uracil) and C (cytosine) were equally distributed (25% each) at the 5′ end of each of the 21- and 24-nt small RNAs. These results indicate that different size classes of small RNAs exhibit different 5′ bias.

miRNA identification in soybean root hairs and stripped root

To identify miRNAs from the three small RNA libraries generated in this study, sequence tags of 18–26 nt in length were compared with the most up-to-date plant miRNA data set (miRBASE V.20, released in 2013). Additionally, the secondary structures of potential miRNA precursor sequences (see Materials and methods) were predicted using DINAMelt software following RNA3.0 parameters (Markham and Zuker, 2005) and criteria established by Meyers *et al.* (2008). Using these criteria, a total of 114 miRNAs, including 92 known miRNA and 22 novel miRNAs, were identified (Table S2). Of the 92 known miRNAs, 30 are members of 17 miRNA families conserved across diverse plant species, seven miRNAs are conserved among other legumes species, and 52 miRNAs were grouped into 41 miRNA families that might be soybean specific (Table S2). miR166 and miR156 were highest in abundance in the conserved miRNA families, whereas miR1507 and miR1511 were the most abundant in legume- and soybean-specific miRNAs, respectively (Table S2). Interestingly, the abundance of some miRNA* was even higher than the corresponding miRNA, such as miR482*. These results support previous reports that miRNA* may also function in gene silencing (Zhang *et al.*, 2011) or it could indicate that the miRNA is misannotated (Jeong *et al.*, 2011).

Identification of *B. japonicum*-responsive soybean miRNAs

To identify tissue-specific *B. japonicum*-responsive miRNAs in soybean, the normalized expression levels of miRNAs in the three libraries (Un_RH, In_RH and stripped roots) were compared. A total of 86 miRNAs were expressed in all three libraries. Six miRNAs appeared specific for root hairs (either infected or not), and two, one and one miRNAs were specifically expressed in the In_RH, Un_RH and stripped root libraries, respectively (Figure 2). Statistical analysis of miRNA abundance between libraries identified 66 miRNAs that were differentially expressed between root hairs and stripped roots (Figure 3a). In addition, we identified 48 miRNAs that were differentially regulated by *B. japonicum* infection through comparison of the In_RH and Un_RH libraries (Figure 3b). To validate these results, the expression levels of 11 *B. japonicum*-responsive soybean miRNAs were measured by quantitative RT-PCR (qRT-PCR) following the protocol described by Varkonyi-Gasic *et al.* (2007). These results are shown in Figure 4. The data generated by qRT-PCR and sequencing were consistent with the up- or down-regulation of genes predicted by the miRNA sequencing results (correlation coefficient of the two data sets is 0.78), although the fold changes were sometimes different. For example, miR1510 was increased 22-fold after *B. japonicum* inoculation based on Illumina sequencing, but only twofold based on miRNA qRT-PCR (Figure 4). This apparent difference may be due to the specificity of miRNA qRT-PCR primer, as opposed to multiple reads sequenced by Illumina.

Identification of miRNA cleavage targets by PARE library sequencing

The small RNA sequences generated from the sequencing of the Un_RH, In_RH and stripped root libraries identified 92 known miRNAs and 22 novel miRNAs. Because biological interpretation of these results is incomplete without identification of the corresponding mRNA targets of the root hair

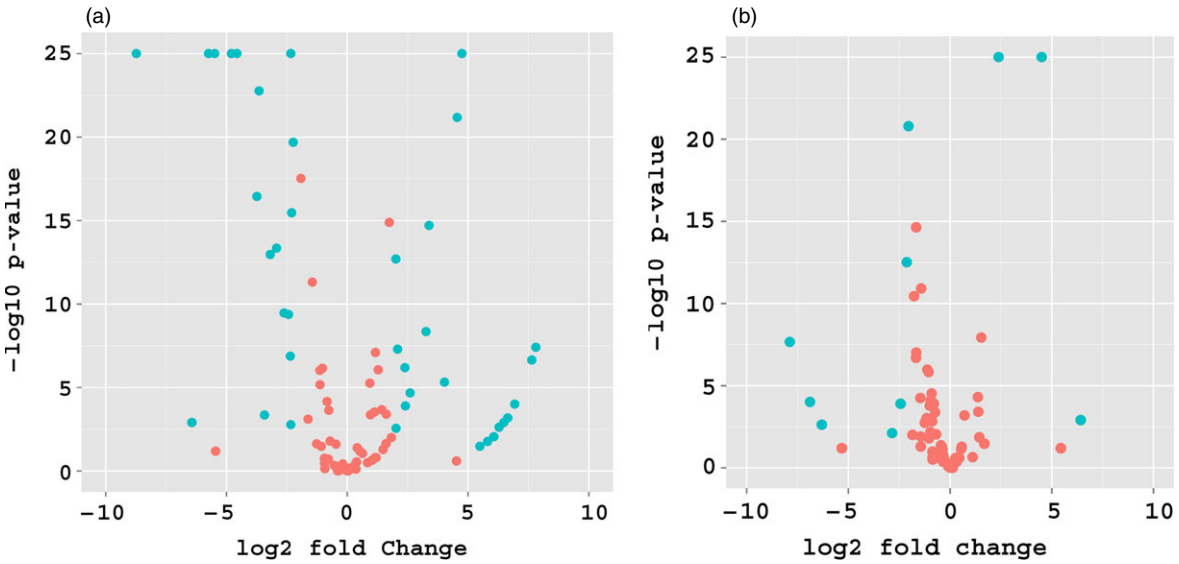

Figure 3 Volcano plot of differential miRNAs expression between libraries. (a) Differential miRNAs expression between root hairs (In_RH) and stripped roots. (b) Differential miRNAs expression between mock-inoculated root hairs (Un_RH) and *Bradyrhizobium japonicum*-infected root hairs (In_RH). Fold change and *P* values between libraries were calculated by edgeR method. *P* values were transformed into negative log10 and miRNAs fold change between libraries was transformed into log2. Negative log10 *P* values and log2 fold change are on the *y*- and *x*-axes, respectively. miRNAs with more than fourfold change and *P* values <0.05 were shown in blue.

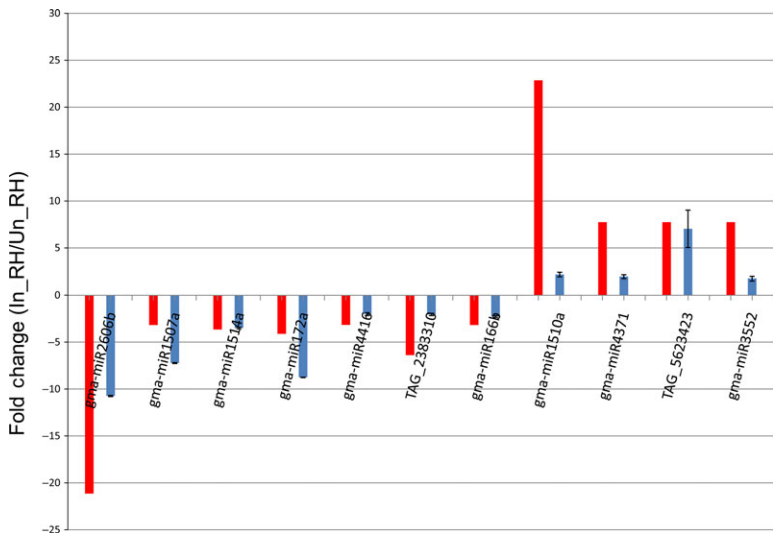

Figure 4 Validation of several rhizobium-responsive miRNA candidates in root hairs by qRT-PCR. The expression levels of 11 selected miRNAs were measured by qRT-PCR (blue column) using cDNAs synthesized from RNA samples collected from In_RH or Un_RH tissues. Three replications were performed. The fold change of each miRNA was calculated by In_RH/Un_RH. The red bar indicates the fold change of each miRNA (In_RH/Un_RH) calculated from Illumina sequencing data.

miRNAs, we constructed a PARE library from the mixture of In_RH and Un_RH RNA. PARE sequencing generated more than 7.9 million raw reads (Table 2), including 83% of 20 or 21 nt sequences, which represents the length of the fragments generated by the *Mme*I enzyme (see Materials and methods). The results of our root hair PARE data were then combined with the previously published soybean seed PARE data (Song *et al.*, 2011), yielding a total of 405 soybean targets (Table S3), which could be cleaved by the miRNAs identified in our study (Figure S1). These target mRNAs were subsequently grouped into different functional categories using the MapMan program (Usadel *et al.*, 2005). MAPMAN is a Web-based tool that classifies genes based on defined categories and displays these predictions graphically (e.g. within the context of metabolic pathways). MapMan analysis showed that most of the identified miRNA targets are predicted to encode proteins involved in transcriptional regulation, protein modification, protein degradation and hormone synthesis (Figure S2).

Table 2 Number of sequenced tags matching the soybean genome, t/r/sn/snoRNA and distinct tags in root hair PARE library

Total sequences	Genome matched	Distinct genome matched	t/r/sn/snoRNA matched
7 943 293	4 020 331	332 989	70 577

gma-miR2606b and gma-miR4416, *B. japonicum*-responsive miRNAs, regulate soybean nodulation

Utilizing small RNA sequence results, we selected four miRNAs, TAG_2383310, miR1514, gma-miR2606b and gma-miR4416, for further functional analysis. The expression levels of these miRNAs were significantly regulated upon *B. japonicum* inoculation (Figure 4). These four *B. japonicum*-responsive miRNAs are legume-(gma-miR2606b and miR1514) or soybean-specific (TAG_2383310 and gma-miR4416) miRNAs.

To investigate the functional roles of these miRNAs in soybean nodulation, we cloned the sequence of the precursor mRNAs for each of the above miRNAs behind the strong CvMV promoter in the binary vector pCAMGFP-CvMV:GWOX, which expresses a constitutive GFP cassette (Li *et al.*, 2010). These constructs were introduced into *Agrobacterium rhizogenes* and subsequently used to transform soybean roots (Collier *et al.*, 2005). The transformed hairy roots were identified by their constitutive GFP expression, and the nodulation phenotypes were recorded for each of the constructs and compared to control roots similarly transformed with an empty vector. As shown in Figure 5a, nodule numbers increased 2.4-fold in hairy roots ectopically expressing gma-miR2606b. In contrast, nodule numbers were decreased by 40% in roots expressing gma-miR4416 (Figure 5a). No significant differences were observed in nodule morphology based on microscopic analysis of nodule structure in thin sections (Figure 5b). No significant changes in nodulation were detected in roots ectopically expressing either miR1514 or TAG_2383310 (Figure S3).

Because ectopic expression of gma-miR2606b and gma-miR4416 altered nodule numbers, we queried our PARE library data to determine whether any of the miRNA targets identified in our study are regulated by gma-miR2606b and gma-miR4416. We identified 2 targets for gma-miR2606b and six targets for gma-miR4416 (Table S3). To validate miRNA activity on these targets, we used qRT-PCR to measure the expression levels of both miRNA and their target mRNAs in tissues taken from the control and miRNA-overexpressing transgenic roots (Figure 6).

For gma-miR2606b, the level of gma-miR2606b was increased more than ninefold relative to control roots with a concomitant reduction in the expression of one of the target (Glyma07g02290) mRNAs (Figure 6a). Glyma07g02290 encodes a *Mannosyl-oligosaccharide 1, 2-alpha-mannosidase* (*MNS*). The role for *MNS* in nodulation is still unknown. However, it was reported that *MNS* might be involve in N-Glycan processing, which affects root development and cell wall biosynthesis (Liebminger *et al.*, 2009).

The level of gma-miR4416 was increased more than 82-fold in the transformed roots with a concomitant reduction in the mRNA abundance of three of the six target genes (Figure 6b). Among these three significantly regulated target genes, Glyma11g29920 was reduced 50-fold in gma-miR4416 overexpression tissues.

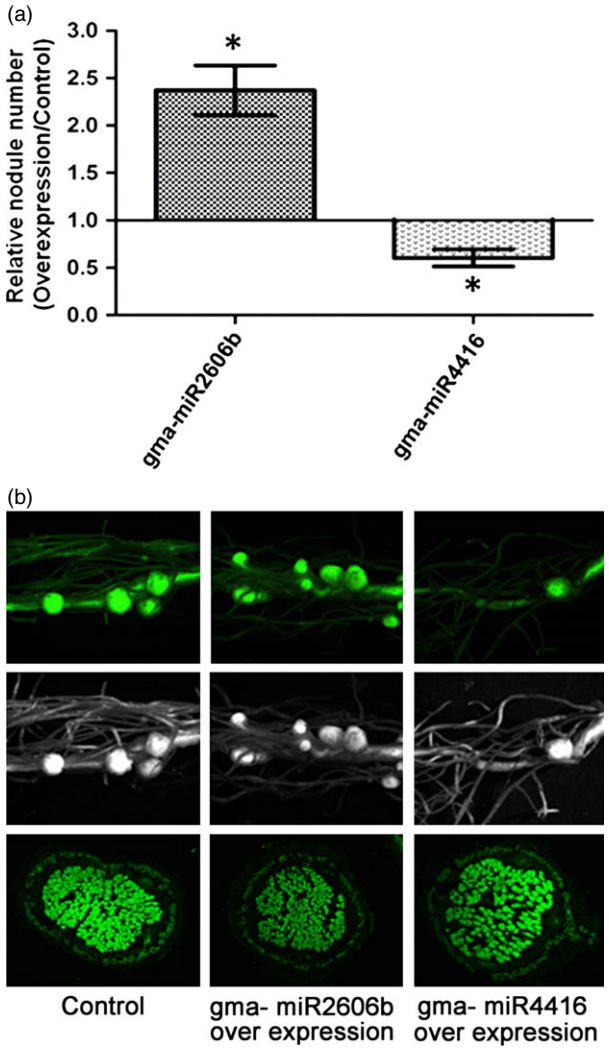

Figure 5 Ectopic expression of miR2606b and miR4416 in soybean hairy roots resulted in significant changes of nodule numbers. (a) Relative nodule number formed on soybean hairy roots expressing miR2606b and miR4416 as compared to roots transformed with the empty vector control. Asterisks indicate that the difference is statistically significant at $P < 0.05$ level. Error bar represents ±SE. (b) GFP-expressing root nodules in transgenic roots (top) and thin section (bottom) of nodule showing the infected and uninfected cells.

Glyma11g29920 encodes a peroxidase gene that is homologous to *M. truncatula RIP1* (*rhizobium-induced peroxidase 1*) with 75.6% similarity in protein sequence (Figure S4) and, hence, we refer to this gene as *GmRIP1*. *RIP1* is a well-studied nodulin gene in *M. truncatula*. It was reported that *RIP1* is induced very early (3 h) by rhizobial infection and is highly expressed in the differentiating root hair during the pre-infection stage in *M. truncatula* roots (Cook *et al.*, 1995). When we performed qRT-PCR to measure *GmRIP1* expression in the Un_RH and In_RH tissues, we found that *GmRIP1* expression was significantly induced after *B. japonicum* infection (Figure 7), similar to the situation in *M. truncatula* (Cook *et al.*, 1995), which is negatively correlated with the expression of gma-miR4416. Taken together, our results suggest that gma-miR4416-mediated regulation of *GmRIP1* is a key regulator during soybean nodulation and symbiosis.

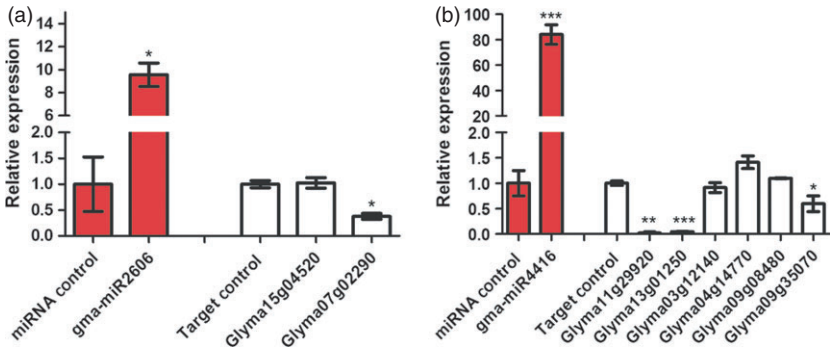

Figure 6 Expression of miRNAs and the corresponding target genes in miRNA overexpression tissues. (a) qRT-PCR analysis of gma-miR2606 (in red bar) and its two targets (in white bar) in gma-miR2606 overexpression tissues. (b) qRT-PCR analysis of gma-miR4416 (in red bar) and its six targets (in white bar) in gma-miR4416 overexpression tissues. Three biological replicates were used. *indicates significant at the 0.05 probability level; ** indicates significant at the 0.01 probability level; *** indicates significant at the 0.005 probability level. Error bar represents ±SE.

Discussion

Although a number of miRNAs were identified in different tissues during soybean nodulation (Joshi *et al.*, 2010; Subramanian *et al.*, 2008; Turner *et al.*, 2012; Wang *et al.*, 2009), there is no report focused on root hairs, which are the primary route of rhizobial infection. High-throughput sequencing identified a total of 114 miRNAs in soybean roots, including 22 novel miRNAs. At the same time, 405 soybean miRNA targets were found through PARE analysis. Notably, the expression levels of 48 miRNAs in root hairs were significantly affected by *B. japonicum* inoculation, an observation that is consistent with the hypothesis that miRNAs play an important regulatory role in early steps of symbiotic development. Of the 48 *B. japonicum*-responsive miRNAs, 38 match known miRNAs and ten miRNAs are novel miRNAs. Among the 38 known miRNAs, 15 are conserved in plant species, seven are conserved in legumes, and 16 are soybean-specific miRNAs. Gene targets of these miRNAs were grouped into functional categories as defined by MapMan (Figure S2). For miRNAs regulated during symbioses, the majority of the targets were predicted to encode transcription factors and, hence, would be expected to mediate their effects by regulating a variety of genes that might be critical to the nodulation process. Indeed, there are a number of reports where targeting of specific transcription factors by miRNAs regulates nodulation; for example, miR169 regulating *HAP2* (Combier *et al.*, 2006), miR166 regulating *HD-ZIPIII* (Boualem *et al.*, 2008) and miR172 regulating *AP2* transcription factors (Wang *et al.*, 2014; Yan *et al.*, 2013). Other targets are also expected to have broad effects on nodulation as they are predicted to play roles in hormone physiology (e.g. IAA, ABA, ethylene and jasmonic acid), or in the perception (e.g. receptor kinases) and transmission (e.g. calcium signalling) of important signal molecules. A few 14-3-3 genes, which encode phosphoserine- or phosphothreonine-binding proteins, were also targeted, suggesting a possible role in modulating the effects of protein phosphorylation.

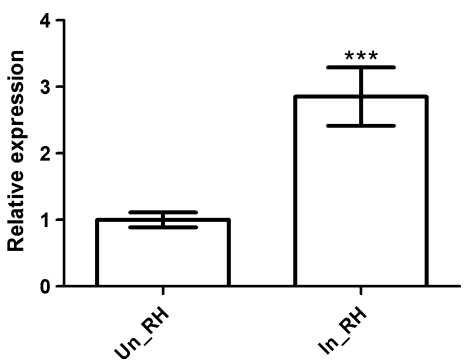

Figure 7 Relative expression levels of *GmRIP1* (Glyma11g29920) in water mock-inoculated control (Un_RH) and *Bradyrhizobium japonicum*-innoculated root hairs (In_RH). Three replications were performed. Bar represents ±SE. *** indicates significant at the 0.005 probability level. Error bar represents ±SE.

miRNA qRT-PCR (Figure 4) was used to validate the expression patterns of specific miRNAs with a focus on those whose expression responded to *B. japonicum* inoculation. Among these miRNAs, miR172 was reported to be a positive regulator of soybean nodulation through targeting several AP2 transcription factors (Wang *et al.*, 2014; Yan *et al.*, 2013). Ectopic overexpression of miR172 increased the number of deformed root hairs and nodule initiation and enhanced nitrogenase activity. Previous research demonstrated that miR166 targets class III *HD-ZIP* transcription factors, *MtCNA1* and *MtCNA2*, and ectopic expression of miR166 reduced lateral root growth and nodule development in *M. truncatula* (Boualem *et al.*, 2008). In our experiment, miR166 levels dropped roughly two-thirds after rhizobial infection (Figure 4). In our PARE data, one *HD-ZIP* transcription factor, Glyma05g30000, was identified as a target of miR166. According to the protein sequence, the Glyma05g30000 coding sequence is about 79.6% and 78.4% similar to the protein sequences of MtCNA1 and MtCNA2, respectively, suggesting that the role of miR166 in nodulation may be conserved between *M. truncatula* and soybean.

miR1507 was identified in soybean root, stem, leaf and mature green seed (Li *et al.*, 2010). Expression analysis found that miR1507 levels did not change in soybean roots from 0 to 14 days after rhizobial infection (Li *et al.*, 2010). However, in our root hair experiments, there was a remarkable decrease in miR1507 levels after rhizobial infection. miR1507 is a 22 nt miRNA and targets a NB-LRR-type disease resistance gene, Glyma04g29220, found in our PARE data. miR1507 was previously reported to target and generate secondary-phased siRNAs from Glyma04g29220, and secondary siRNAs that could regulate

in cis or trans other NB-LRR genes (Arikit *et al.*, 2014). These data suggest that miR1507 could regulate the interaction between root hairs with rhizobium during early infection and might control the early symbiotic process.

While our data clearly identified a variety of interesting genes worthy of further study, we focused on several specific miRNAs, based on their transcriptional response to *B. japonicum*. Ectopic expression of miRNA in soybean roots transformed by *A. rhizogenes* demonstrated that some rhizobium-responsive miRNAs could play important roles during nodulation.

miR2606b is a legume-conserved miRNA. In *M. truncatula*, miR2606b was identified from both roots and nodules (Devers *et al.*, 2011; Lelandais-Briere *et al.*, 2009). The expression level of miR2606b did not change during mycorrhizal infection of roots (Devers *et al.*, 2011). In our experiment, we validated that gma-miR2606b is significantly repressed after *B. japonicum* infection by both illumina sequencing and miRNA qRT-PCR analysis. Overexpression of gma-miR2606b significantly increased soybean nodulation. Target analysis confirmed that gma-miR2606b regulates the expression of a *Mannosyl*-oligosaccharide 1, 2-alpha-mannosidase in soybean roots. Until now, there are no reports describing a role for *MNS* in nodulation. However, in *Arabidopsis*, *MNS* was reported to function in N-linked oligosaccharide processing, which impacts root development and cell wall biosynthesis (Liebminger *et al.*, 2009). In *MNS* mutant plants, root cell wall architecture is altered and plants also display shorter and increased lateral roots (Liebminger *et al.*, 2009). These results suggest that MNS could have a role in root and root hair growth or infection thread formation during the legume microbe symbiosis.

gma-miR4416 is a soybean-specific miRNA of unknown function. According to previous small RNA analysis in different soybean tissues (Joshi *et al.*, 2010), the levels of gma-miR4416 are highest in roots and nodules, but relatively low in seed and flower tissues. In root hairs, the expression of gma-miR4416 was repressed after rhizobial infection. Utilizing our PARE results and confirmation by qRT-PCR, we identified *GmRIP1*, a peroxidase gene that is targeted by gma-miR4416. *RIP1* was reported as an early nodulin gene, which is rapidly induced by rhizobial infection and maintains high expression during the pre-infection period in *M. truncatula* roots (Cook *et al.*, 1995). *RIP1* expression dropped 48 h postinoculation and decreased to levels comparable to uninoculated roots during nodule morphogenesis (Peng *et al.*, 1996). *In situ* hybridization confirmed that *RIP1* was strongly induced in differentiating root hairs, which indicated that *RIP1* could facilitate rhizobial infection (Cook *et al.*, 1995). Further analysis revealed that rhizobial or Nod factor inoculation could induce *RIP1* expression and ROS production with a similar spatial and temporal pattern. *RIP1* expression is also induced by ROS (Ramu *et al.*, 2002). ROS levels increase during the early stages of nodulation elicited by the rhizobial lipo-chitooligosaccharide Nod factor, while this does not occur when plants are inoculated with rhizobial mutant strains that cannot produce this signal (Ramu *et al.*, 2002). ROS is required for root hair outgrowth and initiation of nodulation (D'Haeze *et al.*, 2003). Suppression of ROS accumulation in *M. truncatula* inhibited infection thread development and root hair curling in the rhizobial infection process (Peleg-Grossman *et al.*, 2007). In soybean, gma-miR4416 expression is root hair specific, relative to the other tissues examined, such as root tip, root, nodule, apical meristem, leaf, flower and green pod (Libault *et al.*, 2010b). In our research, we found that *GmRIP1* expression negatively correlates with gma-miR4416 expression levels and increased after rhizobial infection in soybean root hairs. The data are consistent with a model by which gma-miR4416 targets *GmRIP1*, which could potentially mediate ROS levels such as to promote early infection but prohibit any later, negative effects. Although the function of *MNS* (targeted by miR2606b) in legume nodulation and infection is not known, it may have role in root or root hair growth and the infection process. On the other hand, consistent with the previous work in *M. truncatula*, *RIP1* (targeted by gma-miR4416) clearly has an important role to play during the earliest stages of the infection process.

Materials and methods

Plant materials and library construction

Root hair and stripped root tissues were isolated from soybean [*G. max* (L) Merr.] cultivar Williams 82, as described by Brechenmacher *et al.* (2010). Briefly, seeds were surface sterilized with 30% bleach, rinsed several times with autoclaved double-distillated H_2O (ddH_2O), once with 0.01 M HCl for 10 min, and then several times with autoclaved ddH_2O. Sterilized seeds were put in plates contained nitrogen-free B&D agar medium (Broughton and Dilworth, 1971) and germinated in a growth chamber in the dark at 27 °C and 80% humidity. After 3 days, soybean roots were inoculated with *B. japonicum* strain USDA110 or mock-inoculated with sterile water. Root hairs and stripped roots were separated after freezing in liquid nitrogen.

Tissues, uninoculated root hairs (Un-RH), inoculated root hairs (In_RH) or stripped roots, were collected 12, 18, 24 and 48 h postinoculation. Each time point had three biological replicates. Total RNA was purified using Trizol Reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, CA) and subsequently purified using chloroform extraction. Three small RNAs libraries (uninoculated root hairs, inoculated root hairs or stripped roots) were generated. For each library, 12 RNA samples from three biological replicates and four time points were pooled equally. Mixed RNA was used for small RNA library construction following the protocol described by Lu *et al.* (2007). Briefly, small molecular weight RNA was separated from total RNA by precipitation in the presence of 5 M NaCl and PEG8000. The 18–30 nt small RNA size class was purified from a 15% polyacrylamide/urea gel. Purified small RNAs were ligated to 3′ and 5′ adapters. After reverse transcription and PCR amplification, enriched small RNAs (sRNAs) were sequenced using the Illumina HiSeq 2000 platform at the University of Delaware's Sequencing and Genotyping Center.

Small RNA library analysis

Small RNA analyses were performed as described by Joshi *et al.* (2010). Briefly, small RNA sequences were trimmed of the adapter sequence and mapped to the soybean genome. Only perfectly matched sequences were retained for analysis. The sequences that matched soybean tRNA and rRNA sequences were excluded from further analysis. Small RNA abundance was normalized into transcripts per million (TPM) to compare small RNA expression between libraries.

miRNA prediction and statistical analysis

The method developed by Joshi *et al.* (2010) was used to identify miRNAs from the soybean small RNA sequences. Briefly, the predicted miRNA precursor sequence including the 200 bp

upstream and downstream sequences was extracted from the published soybean genome sequence and the derived RNA secondary structures predicted using the DINAMelt program with default RNA3.0 parameters (Markham and Zuker, 2005). The criteria defined by Meyers *et al.* (2008) were used to distinguish false and positive miRNA secondary structures predicted by DINAMelt.

To compare the abundance of miRNAs between libraries, the edgeR (Robinson and Smyth, 2008; Robinson *et al.*, 2010) method was applied to analyse the data using the method for data without biological replicates. For controlled experiments, 0.1 is suggested in the user guide (Chen *et al.*, 2014) as the square-root-dispersion parameter for genetically identical model organisms. An exact negative binomial test was used to test each miRNA for differential expression and *P*-values were computed. False discovery rates (FDRs) were then calculated from the list of *P*-values using the BH algorithm (Benjamini and Hochberg, 1995). We chose 0.05 as the cut-off for controlling FDR.

miRNA target analysis

A soybean PARE library was generated according to the methods described by German *et al.* (2009). To determine potential targets of miRNAs identified in the inoculated and uninoculated soybean root hairs, as well as in the stripped roots, we used the pipelines which were well described by Zhai *et al.* (2011). Mature miRNA sequences were used as queries to search for potential target messenger RNA (mRNA). To confirm the potential target mRNA generated by the CleaveLand pipeline (Addo-Quaye *et al.*, 2009). We mined a previously published soybean degradome library developed from soybean seed tissue (Song *et al.*, 2011), which contains 15 million degradome sequencing reads, including about one million distinct genome-matched reads. For degradome analysis, two windows flanking each predicted cleavage site were defined. The small window (Ws) was defined as 2 nt around the cleavage site, and a large window (Wl) was defined as 15 nt around cleavage site. The sum of the abundance of tags located in each window region was calculated. Targets were accepted if Ws $\geq$ 2 with a corresponding Ws/Wl $\geq$ 0.5.

qRT-PCR analysis of microRNAs and mRNA targets

Quantification of mature miRNA abundance was performed by quantitative RT-PCR (qRT-PCR) following the protocol described by Varkonyi-Gasic *et al.* (2007). Briefly, genomic DNA (gDNA) was removed from the purified RNA using Turbo DNase (Ambion Inc., Austin, TX) following the manufacturer's instructions. One microgram of gDNA-free RNA was reverse-transcribed using Superscript III (Invitrogen) using a stem-loop-specific RT-primer, which contained six nucleotide-specific sequences from the complementary 3′ end sequence of each miRNA (Table S4). qRT-PCR was performed using SYBR Green PCR mix (ABI enzyme) using a miRNA-specific forward primer and a universal reverse primer. The reference genes *snoR1* and 5.8S (Table S4) were selected to calculate the relative expression of each miRNA.

We analysed the expression of target mRNA by quantitative RT-PCR following the protocol previously described using *Con4* and *Con15* as reference genes (Libault *et al.*, 2008) and relative expression data were analysed according to Libault *et al.* (2010b).

Three replications were performed for both miRNAs and mRNA targets analyses, and Student's *t*-test used to compare differences between control and experimental values.

Plasmid construction and soybean hairy root transformation

Agrobacterium rhizogenes-dependent hairy root transformation was used to construct transgenic soybean roots ectopically expressing the selected miRNAs. The vector pCAMGFP-DC-GWOX was used, which allowed cloning of the precursor mRNA of each of the miRNA behind the strong CvMV promoter. The T-DNA of this vector also contained a constitutively expressed GFP cassette (Li *et al.*, 2010), which enabled identification of transformed roots by GFP fluorescence. The nodulation phenotypes of only the transformed roots were measured and compared to empty vector controls.

To clone miRNA precursors, the region of DNA spanning 200 bp upstream and downstream of each miRNA coding region was amplified from soybean genomic DNA (Table S4). The amplified precursors were cloned in the pDONR-Zeo vector using the Gateway® BP Clonase® II enzyme mixes (Invitrogen) and sequenced. The resulting plasmids were recombined into the pCAMGFP-DC-GWOX binary vector using the Gateway® LR Clonase® II enzyme mixes (Invitrogen). Subsequently, the positive plasmids were introduced into *A. rhizogenes* K599 by electroporation and used for hairy root transformation as previously described (Libault *et al.*, 2009). Briefly, apical stem sections were excised from 14-day-old soybean seedlings and infected by *A. rhizogenes* K599 harbouring either empty vector as a control or the miRNA precursor overexpression constructs. After 24 days of incubation, the soybean hairy roots were inoculated with *B. japonicum* USDA110. Transgenic roots were identified 6 weeks postinoculation based on their strong GFP fluorescence. Hairy roots were generated on 12 plants per construct, and the transformation repeated six times under the same conditions. For each replicate, the fold change of nodule numbers of miRNA-expressing roots compared to control root was calculated to evaluate the effect of each miRNA on nodulation. To test whether there was a significant fold change of nodule number after miRNA overexpression, a hypothesis *t*-test was performed with the null hypothesis of fold change of nodule number equals 1. For nodule morphology analysis, nodule hand sections were stained with SYTO13 dye (Invitrogen) described by Haynes *et al.* (2004). For nodule ultrastructure, nodules were embedded in paraplast and sectioned prior to staining with toluidine blue, paraplast was removed, and then, nodule sections were visualized by light microscopy.

Acknowledgements

Research was funded by a grant from the USA National Science Foundation Plant Genome program (Grant No. DBI-0421620 to GS and DX).

References

Addo-Quaye, C., Miller, W. and Axtell, M.J. (2009) CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. *Bioinformatics*, **25**, 130–131.

Allen, E., Xie, Z., Gustafson, A.M. and Carrington, J.C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, **121**, 207–221.

Arikit, S., Xia, R., Kakrana, A., Huang, K., Zhai, J., Yan, Z., Valdes-Lopez, O., Prince, S., Musket, T.A., Nguyen, H.T., Stacey, G. and Meyers, B. (2014) An

atlas of soybean small RNAs identifies phased siRNAs from hundreds of coding genes. *Plant Cell*, **26**, 4584–4601.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B*, **57**, 289–300.

Boualem, A., Laporte, P., Jovanovic, M., Laffont, C., Plet, J., Combier, J.P., Niebel, A., Crespi, M. and Frugier, F. (2008) MicroRNA166 controls root and nodule development in *Medicago truncatula*. *Plant J.* **54**, 876–887.

Brechenmacher, L., Lei, Z., Libault, M., Findley, S., Sugawara, M., Sadowsky, M.J., Sumner, L.W. and Stacey, G. (2010) Soybean metabolites regulated in root hairs in response to the symbiotic bacterium *Bradyrhizobium japonicum*. *Plant Physiol.* **153**, 1808–1822.

Brechenmacher, L., Nguyen, T.H., Hixson, K., Libault, M., Aldrich, J., Pasa-Tolic, L., Stacey, G. (2012) Identification of soybean proteins from a single cell type: the root hair. *Proteomics*, **12**, 3365–3373.

Broughton, W.J. and Dilworth, M.J. (1971) Control of leghaemoglobin synthesis in snake beans. *Biochem. J.* **125**, 1075–1080.

Chen, X. (2009) Small RNAs and their roles in plant development. *Annu. Rev. Cell Dev. Biol.* **25**, 21–44.

Chen, X. (2012) Small RNAs in development—insights from plants. *Curr. Opin. Genet. Dev.* **22**, 361–367.

Chen, Y., McCarthy, D., Robinson, M.D. and Smyth, G.K. (2014) *edgeR: differential expression analysis of digital gene expression data User's Guide*. www.bioconductor.org

Collier, R., Fuchs, B., Walter, N., Kevin Lutke, W. and Taylor, C.G. (2005) Ex vitro composite plants: an inexpensive, rapid method for root biology. *Plant J.* **43**, 449–457.

Combier, J.P., Frugier, F., de Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., Vernie, T., Ott, T., Gamas, P., Crespi, M. and Niebel, A. (2006) MtHAP2-1 is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev.* **20**, 3084–3088.

Cook, D., Dreyer, D., Bonnet, D., Howell, M., Nony, E. and VandenBosch, K. (1995) Transient induction of a peroxidase gene in *Medicago truncatula* precedes infection by *Rhizobium meliloti*. *Plant Cell*, **7**, 43–55.

Devers, E.A., Branscheid, A., May, P. and Krajinski, F. (2011) Stars and symbiosis: microRNA- and microRNA*-mediated transcript cleavage involved in arbuscular mycorrhizal symbiosis. *Plant Physiol.* **156**, 1990–2010.

D'Haeze, W., De Rycke, R., Mathis, R., Goormachtig, S., Pagnotta, S., Verplancke, C., Capoen, W. and Holsters, M. (2003) Reactive oxygen species and ethylene play a positive role in lateral root base nodulation of a semiaquatic legume. *Proc. Natl Acad. Sci. USA*, **100**, 11789–11794.

Eddy, S.R. (2001) Non-coding RNA genes and the modern RNA world. *Nat. Rev. Genet.* **2**, 919–929.

German, M.A., Luo, S., Schroth, G., Meyers, B.C. and Green, P.J. (2009) Construction of parallel analysis of RNA ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. *Nat. Protoc.* **4**, 356–362.

Haynes, J.G., Czymmek, K.J., Carlson, C.A., Veereshlingam, H., Dickstein, R. and Sherrier, D.J. (2004) Rapid analysis of legume root nodule development using confocal microscopy. *New Phytol.* **163**, 661–668.

Hewezi, T., Maier, T.R., Nettleton, D. and Baum, T.J. (2012) The arabidopsis microRNA396-GRF1/GRF3 regulatory module acts as a developmental regulator in the reprogramming of root cells during cyst nematode infection. *Plant Physiol.* **159**, 321–335.

Jeong, D.H., Park, S., Zhai, J., Gurazada, S.G., De Paoli, E., Meyers, B.C. and Green, P.J. (2011) Massive analysis of rice small RNAs: mechanistic implications of regulated microRNAs and variants for differential target RNA cleavage. *Plant Cell*, **23**, 4185–4207.

Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E. and Walker, G.C. (2007) How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model. *Nat. Rev. Microbiol.* **5**, 619–633.

Joshi, T., Yan, Z., Libault, M., Jeong, D.H., Park, S., Green, P.J., Sherrier, D.J., Farmer, A., May, G., Meyers, B.C., Xu, D. and Stacey, G. (2010) Prediction of novel miRNAs and associated target genes in *Glycine max*. *BMC Bioinformatics*, **11**(Suppl. 1), S14.

Kasschau, K.D., Fahlgren, N., Chapman, E.J., Sullivan, C.M., Cumbie, J.S., Givan, S.A. and Carrington, J.C. (2007) Genome-wide profiling and analysis of Arabidopsis siRNAs. *PLoS Biol.* **5**, e57.

Katiyar-Agarwal, S. and Jin, H. (2010) Role of small RNAs in host-microbe interactions. *Annu. Rev. Phytopathol.* **48**, 225–246.

Khraiwesh, B., Zhu, J.K. and Zhu, J. (2012) Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochim. Biophys. Acta*, **1819**, 137–148.

Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993) The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, **75**, 843–854.

Lelandais-Briere, C., Naya, L., Sallet, E., Calenge, F., Frugier, F., Hartmann, C., Gouzy, J. and Crespi, M. (2009) Genome-wide *Medicago truncatula* small RNA analysis revealed novel microRNAs and isoforms differentially regulated in roots and nodules. *Plant Cell*, **21**, 2780–2796.

Li, H., Deng, Y., Wu, T., Subramanian, S. and Yu, O. (2010) Misexpression of miR482, miR1512, and miR1515 increases soybean nodulation. *Plant Physiol.* **153**, 1759–1770.

Libault, M., Thibivilliers, S., Bilgin, D.D., Radwan, O., Benitez, M., Clough, S.J. and Stacey, G. (2008) Identification of four soybean reference genes for gene expression normalization. *Plant Genome*, **1**, 44–54.

Libault, M., Joshi, T., Takahashi, K., Hurley-Sommer, A., Puricelli, K., Blake, S., Finger, R.E., Taylor, C.G., Xu, D., Nguyen, H.T. and Stacey, G. (2009) Large-scale analysis of putative soybean regulatory gene expression identifies a Myb gene involved in soybean nodule development. *Plant Physiol.* **151**, 1207–1220.

Libault, M., Brechenmacher, L., Cheng, J., Xu, D. and Stacey, G. (2010a) Root hair systems biology. *Trends Plant Sci.* **15**, 641–650.

Libault, M., Farmer, A., Joshi, T., Takahashi, K., Langley, R.J., Franklin, L.D., He, J., Xu, D., May, G. and Stacey, G. (2010b) An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants. *Plant J.* **63**, 86–99.

Liebminger, E., Huttner, S., Vavra, U., Fischl, R., Schoberer, J., Grass, J., Blaukopf, C., Seifert, G.J., Altmann, F. and Mach, L. (2009) Class I alpha-mannosidases are required for N-glycan processing and root development in Arabidopsis thaliana. *Plant Cell*, **21**, 3850–3867.

Lu, C., Meyers, B.C. and Green, P.J. (2007) Construction of small RNA cDNA libraries for deep sequencing. *Methods*, **43**, 110–117.

Markham, N.R. and Zuker, M. (2005) DINAMelt web server for nucleic acid melting prediction. *Nucleic Acids Res.* **33**, W577–W581.

Meyers, B.C., Axtell, M.J., Bartel, B., Bartel, D.P., Baulcombe, D., Bowman, J.L., Cao, X., Carrington, J.C., Chen, X., Green, P.J., Griffiths-Jones, S., Jacobsen, S.E., Mallory, A.C., Martienssen, R.A., Poethig, R.S., Qi, Y., Vaucheret, H., Voinnet, O., Watanabe, Y., Weigel, D. and Zhu, J.K. (2008) Criteria for annotation of plant MicroRNAs. *Plant Cell*, **20**, 3186–3190.

Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., Chen, S., Hannon, G.J. and Qi, Y. (2008) Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5′ terminal nucleotide. *Cell*, **133**, 116–127.

Nguyen, T.H., Brechenmacher, L., Aldrich, J.T., Clauss, T.R., Gritsenko, M.A., Hixson, K.K., Libault, M., Tanaka, K., Yang, F., Yao, Q., Pasa-Tolić, L., Xu, D., Nguyen, H.T. and Stacey, G. (2012) Quantitative phosphoproteomic analysis of soybean root hairs inoculated with *Bradyrhizobium japonicum*. *Mol. Cell Proteomics*, **11**, 1140–1155.

Oldroyd, G.E. (2013) Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* **11**, 252–263.

Oldroyd, G.E., Downie, J.A. (2004) Calcium, kinases and nodulation signalling in legumes. *Nat. Rev. Mol. Cell Biol.* **5**, 566–576.

Peleg-Grossman, S., Volpin, H. and Levine, A. (2007) Root hair curling and Rhizobium infection in *Medicago truncatula* are mediated by phosphatidylinositide-regulated endocytosis and reactive oxygen species. *J. Exp. Bot.* **58**, 1637–1649.

Ramu, S.K., Peng, H.M. and Cook, D.R. (2002) Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene *rip1* in *Medicago truncatula*. *Mol. Plant Microbe Interact.* **15**, 522–528.

Robinson, M.D. and Smyth, G.K. (2008) Small sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics*, **9**, 321–332.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

Ruiz-Ferrer, V. and Voinnet, O. (2009) Roles of plant small RNAs in biotic stress responses. *Annu. Rev. Plant Biol.* **60**, 485–510.

Song, Q.X., Liu, Y.F., Hu, X.Y., Zhang, W.K., Ma, B., Chen, S.Y. and Zhang, J.S. (2011) Identification of miRNAs and their target genes in developing soybean seeds by deep sequencing. *BMC Plant Biol.* **11**, 5.

Staiger, D., Korneli, C., Lummer, M. and Navarro, L. (2013) Emerging role for RNA-based regulation in plant immunity. *New Phytol.* **197**, 394–404.

Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W.B., Zhu, J.K. and Yu, O. (2008) Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genom.* **9**, 160.

Turner, M., Yu, O. and Subramanian, S. (2012) Genome organization and characteristics of soybean microRNAs. *BMC Genom.* **13**, 169.

Turner, M., Nizampatnam, N.R., Baron, M., Coppin, S., Damodaran, S., Adhikari, S., Arunachalam, S.P., Yu, O. and Subramanian, S. (2013) Ectopic expression of miR160 results in auxin hypersensitivity, cytokinin hyposensitivity, and inhibition of symbiotic nodule development in soybean. *Plant Physiol.* **162**, 2042–2055.

Usadel, B., Nagel, A., Thimm, O., Redestig, H., Blaesing, O.E., Palacios-Rojas, N., Selbig, J., Hannemann, J., Piques, M.C., Steinhauser, D., Scheible, W.R., Gibon, Y., Morcuende, R., Weicht, D., Meyer, S. and Stitt, M. (2005) Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiol.* **138**, 1195–1204.

Varkonyi-Gasic, E., Wu, R., Wood, M., Walton, E.F. and Hellens, R.P. (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods*, **3**, 12.

Wang, Y., Li, P., Cao, X., Wang, X., Zhang, A. and Li, X. (2009) Identification and expression analysis of miRNAs from nitrogen-fixing soybean nodules. *Biochem. Biophys. Res. Commun.* **378**, 799–803.

Wang, Y., Wang, L., Zou, Y., Chen, L., Cai, Z., Zhang, S., Zhao, F., Tian, Y., Jiang, Q., Ferguson, B.J., Gresshoff, P.M. and Li, X. (2014) Soybean miR172c targets the repressive AP2 transcription factor NNC1 to activate ENOD40 expression and regulate nodule initiation. *Plant Cell*, **26** (12), 4782–4801.

Yan, Z., Hossain, M.S., Wang, J., Valdes-Lopez, O., Liang, Y., Libault, M., Qiu, L. and Stacey, G. (2013) miR172 regulates soybean nodulation. *Mol. Plant Microbe Interact.* **26**, 1371–1377.

Zhai, J., Jeong, D.H., De Paoli, E., Park, S., Rosen, B.D., Li, Y., Gonzalez, A.J., Yan, Z., Kitto, S.L., Grusak, M.A., Jackson, S.A., Stacey, G., Cook, D.R., Green, P.J., Sherrier, D.J. and Meyers, B.C. (2011) MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased, trans-acting siRNAs. *Genes Dev.* **25**, 2540–2553.

Zhang, X., Zhao, H., Gao, S., Wang, W.C., Katiyar-Agarwal, S., Huang, H.D., Raikhel, N. and Jin, H. (2011) Arabidopsis Argonaute 2 regulates innate immunity via miRNA393(*)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. *Mol. Cell*, **42**, 356–366.

Zhu, Q.-H., Fan, L., Liu, Y., Xu, H., Llewellyn, D. and Wilson, I. (2013) miR482 Regulation of NBS-LRR defense genes during fungal pathogen infection in cotton. *PLoS One*, **8**, e84390.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Example of two miRNA targets validated by PARE analysis.
Figure S2 Distribution of soybean miRNA targets using MapMan categories.
Figure S3 Relative nodule number formed on soybean hairy roots expressing miR1514 and TAG_2383310 as compared to the control roots transformed with the empty vector.
Figure S4 Alignment of amino acid sequences of *GmRIP1* with *MtRIP1*.
Table S1 Distribution of identity of the first nucleotide in different length of small RNA.
Table S2 Summary of miRNA candidates identified in soybean.
Table S3 miRNA targets validated by PARE analysis.
Table S4 primers used in this study.

