frontiers in PLANT SCIENCE

IMPACT 3.6 FACTOR

Plant Genetics and Genomics

Genome-wide identification of non-coding RNAs interacted with microRNAs in soybean

Chuyu Ye, Hao Xu, Enhui Shen, Yang Liu, Yu Wang, Yifei Shen, Jie Qiu, Qian-Hao Zhu and Longjiang Fan

Journal Name:	Frontiers in Plant Science
ISSN:	1664-462X
Article type:	Original Research Article
Received on:	24 Sep 2014
Accepted on:	05 Dec 2014
Provisional PDF published on:	05 Dec 2014
www.frontiersin.org:	www.frontiersin.org
Citation:	Ye C, Xu H, Shen E, Liu Y, Wang Y, Shen Y, Qiu J, Zhu Q and Fan L(2014) Genome-wide identification of non-coding RNAs interacted with microRNAs in soybean. *Front. Plant Sci.* 5:743. doi:10.3389/fpls.2014.00743

Genome-wide identification of non-coding RNAs interacted with microRNAs in soybean

Chu-Yu Ye[1], Hao Xu[2], Enhui Shen[1], Yang Liu[1], Yu Wang[1], Yifei Shen[1], Jie Qiu[1], Qian-Hao Zhu[3], Longjiang Fan[1,*]

[1] Institute of Crop Sciences & Institute of Bioinformatics, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China

[2] Guhe information, Hangzhou 310058, China

[3] CSIRO Agriculture Flagship, Canberra, ACT 2601, Australia

[*]Author for correspondence:

Longjiang Fan

Institute of Crop Sciences, College of Agriculture and Biotechnology, Zhejiang University

866 Yuhangtang Rd, Hangzhou 310058, China

Email: fanlj@zju.edu.cn

Running title: ncRNAs interacted with miRNAs

Number of words: 6,911; Figures: 5; Table: 1; Supplementary tables and figure: 9

Abstract

A wide range of RNA species interacting with microRNAs (miRNAs) form a complex gene regulation network and play vital roles in diverse biological processes. In this study, we performed a genome-wide identification of endogenous target mimics (eTMs) for miRNAs and phased-siRNA-producing loci (*PHAS*) in soybean with a focus on those involved in lipid metabolism. The results showed that a large number of eTMs and *PHAS* genes could be found in soybean. Additionally, we found that lipid metabolism related genes were potentially regulated by 28 miRNAs, and nine of them were potentially further regulated by a number of eTMs with expression evidence. Thirty-three miRNAs were found to trigger production of phasiRNAs from 49 *PHAS* genes, which were able to target lipid metabolism related genes. Degradome data supported miRNA- and/or phasiRNA-mediated cleavage of genes involved in lipid metabolism. Most eTMs for miRNAs involved in lipid metabolism and phasiRNAs targeting lipid metabolism related genes showed a tissue-specific expression pattern. Our bioinformatical evidences suggested that lipid metabolism in soybean is potentially regulated by a complex non-coding network, including miRNAs, eTMs and phasiRNAs, and the results extended our knowledge on functions of non-coding RNAs.

Keywords: miRNA, endogenous target mimics, phasiRNA, lipid metabolism, soybean

1. Introduction

Soybean (*Glycine max*) is one of the most important crops in the world. One of its main characteristics is its capacity to fix atmospheric nitrogen through symbioses with microorganisms (Schmutz et al., 2010). Soybean seeds contain abundant protein and oil that are crucial raw materials for food, feed, and other industrial applications (Lardizabal et al., 2008). Many microRNAs (miRNAs) with a potential role in stress responses (Zeng et al., 2010;Kulcheski et al., 2011;Li et al., 2011;Zeng et al., 2012), nodulation (Subramanian et al., 2008;Wang et al., 2009;Li et al., 2010;Turner et al., 2012), and development (Joshi et al., 2010;Song et al., 2011;Wong et al., 2011;Shamimuzzaman and Vodkin, 2012) have been identified in soybean. miRNAs have also been implicated to play key roles in lipid metabolism of oil crops. For example, some miRNAs (e.g., miR156 and miR6029) from *Brassica napus*, another important edible oil crop, were differentially expressed in cultivars with different seed oil content or at different embryonic developmental stages (Zhao et al., 2012). Several studies also identified differentially expressed miRNAs in developing seeds of soybean, suggesting that these miRNAs might be involved in lipid metabolism (Song et al., 2011;Shamimuzzaman and Vodkin, 2012).

Small RNAs, including miRNA, repeat-associated small interfering RNA (ra-siRNA), *trans*-acting siRNA (tasiRNA) and natural antisense siRNA (nat-siRNA), play vital roles in plant development as well as in adaption to biotic and abiotic stresses (Jones-Rhoades et al., 2006;Rubio-Somoza and Weigel, 2011;Khraiwesh et al., 2012). miRNA regulates gene expression by mediating gene silencing at post-transcriptional level. miRNA is processed from primary miRNA (pri-miRNA) generated by RNA polymerase II. In plants, pri-miRNA is cleaved into miRNA precursor (pre-miRNA) containing a hairpin-like structure, which is further cleaved to give rise to a miRNA/miRNA* duplex that is methylated at the 3' ends. miRNA* is generally degraded and the mature miRNA molecule is incorporated into a RNA-induced silencing complex to target complementary mRNAs through either cleavage or translational inhibition (Mallory and Vaucheret, 2006;Banks et al., 2012;Turner et al., 2012). In addition to silencing protein coding mRNAs, miRNAs are able to target *trans*-acting siRNA transcripts (*TAS*) to trigger production of phased tasiRNAs. On the other hand, the activity of miRNAs can be attenuated or abolished by endogenous target mimics (eTMs), which are usually non-coding transcripts and are able to sequester miRNAs in an uncleavable manner (Banks et al., 2012).

The first eTM identified in plants is the non-protein coding gene *INDUCED BY PHOSPHATE STARVATION1* (*IPS1*) from *Arabidopsis thaliana* (Franco-Zorrilla et al., 2007). *IPS1* binds to miR399, the phosphate starvation-induced miRNA. Because of a small loop caused by a few base pairs of mismatches at the expected miRNA cleavage site, *IPS1* RNA is not cleaved by miR399 but instead serving as a decoy for miR399 to interfere binding of miR399 to its canonical target, *PHO2*. Thus, plants overexpressing *IPS1* showed increased accumulation of the expression of miR-399 target *PHO2* and, concomitantly, reduced shoot Pi content. This mechanism of inhibition of miRNA activity is termed as target mimicry (Franco-Zorrilla et al., 2007). Several artificial target mimics (TMs) designed for different miRNAs with a similar paring pattern as that of

IPS1 and miR399 have been proven to affect the functions of their corresponding miRNAs in transgenic plants (Todesco et al., 2010;Ivashuta et al., 2011;Yan et al., 2012). Recently, eTMs of some conserved miRNAs have been computationally identified from intergenic regions or sequences with only short open reading frames in *A. thaliana*, *Oryza sativa* and other plants (Banks et al., 2012;Wu et al., 2013). Additionally, *A. thaliana* transgenic plants overexpressing eTMs of miR160 and miR166 showed altered plant development, suggesting that these eTMs could be endogenously functional (Wu et al., 2013).

Biogenesis of tasiRNAs is triggered by interaction of miRNA at single or dual sites of the non-coding precursor *TAS* transcripts (Allen et al., 2005). Generation of tasiRNAs from *TAS* transcripts shows a phased pattern in which small RNAs are generated precisely in a head-to-tail arrangement starting from the miRNA cleavage site (Fei et al., 2013). It was found that many miRNAs could trigger the production of tasiRNAs, such as miR390, miR173, and miR828 in *A. thaliana* (Allen et al., 2005;Yoshikawa et al., 2005;Axtell et al., 2006;Rajagopalan et al., 2006). miR390 targets *TAS3* at two complementary sites to initiate production of tasiRNAs. miR173 targets both *TAS1* and *TAS2*, and miR828 directs tasiRNA biogenesis from *TAS4*. *TAS3* has been identified across a broad range of species (including moss, gymnosperms, and angiosperms), while *TAS1*, *TAS2*, and *TAS4* are found only in Arabidopsis and its close relatives (Fei et al., 2013;Hu et al., 2013). tasiRNAs mainly regulate the expression levels of their target transcripts in *trans*, e.g., *TAS3*-derived tasiARF targets members of the auxin response factor (ARF) family, including *ARF2*, *ARF3*, and *ARF4* (Williams et al., 2005). In addition to non-coding transcripts, miRNAs have also been shown to trigger production of phased siRNAs (phasiRNAs) from protein-coding loci, so called phasiRNA producing loci or *PHAS* loci, such as genes encoding pentatricopeptide repeat-containing proteins (PPRs), nucleotide-binding and leucine-rich-repeat-containing proteins (NB-LRRs), and MYB transcription factors in *Arabidopsis*, *Medicago*, *Malus*, *Prunus*, and *Nicotiana* (Howell et al., 2007;Zhai et al., 2011;Zhu et al., 2012;Xiao et al., 2014).

miRNAs interact with a diverse RNA species, such as protein coding target mRNAs and *PHAS* genes as well as eTMs, suggesting the presence of a complex gene regulation network involving miRNAs. In this study, we bioinformatically identified eTMs for miRNAs and *PHAS* genes targeted by miRNAs in soybean genome, and investigated their expression profiles in a wide range of organs and tissues. Our results suggest that a complex network including miRNAs, eTMs for miRNAs and phasiRNAs exists in soybean and it play important roles in diverse biological processes including lipid metabolism.

2. Materials and methods

2.1 Genomic, transcriptomic and degradomic data sources

Genome sequences of *G. max* (version 1.0) and other four oil crops (or their progenitor species), *Ricinus communis* (v0.1), *Linum usitatissimum* (v1.0), *Brassica rapa* (v1.2), and *Gossypium raimondii* (v2.1), were obtained from Phytozome (http://www.phytozome.net/;

v9.1) (Goodstein et al., 2012). Intergenic sequences of *G. max* were retrieved according to the information provided in the GFF file (Schmutz et al., 2010). Transcriptomic data (mRNA and small RNA) and degradomic data of *G. max* were downloaded from NCBI with the accession numbers listed in Supplementary Table S1 (Libault et al., 2010;Song et al., 2011;Shamimuzzaman and Vodkin, 2012;Collakova et al., 2013;Hu et al., 2013;Song et al., 2013). *G. max* miRNAs were downloaded from miRBase (Release 20; http://www.mirbase.org/), and those with ≤ 4 and > 4 mismatches compared to miRNAs of other plants in miRBase were considered as conserved miRNAs and soybean-specific miRNAs, respectively (Meyers et al., 2008).

2.2 Identification of eTMs

Intergenic sequences of the *G. max* genome were collected as the eTM prediction library. eTM identification was performed using the bioinformatic pipeline previously described based on the following rules: (1) bulges composed of three nucleotides are only permitted at the positions corresponding to the 9th to 12th nucleotides counting from the 5' end of a miRNA sequence; (2) perfect pairing is required from the 2nd to 8th positions at the 5' end of a miRNA sequence; (3) the total number of mismatches and G/U pairs within the eTM and miRNA pairing region (excluding the central bulge) should be no more than three; and (4) the distance between an eTM and its upstream/downstream genes is longer than 200 nucleotides (Wu et al., 2013).

2.3 Expression and conservation analysis of eTMs

To investigate the expression profiles of the predicted eTMs, publicly available RNA-Seq data generated from various tissues and organs of soybean (Table S1) were mapped to eTM sequences (including the miRNA binding sites and their 50-bp flanking regions) using TopHat (http://tophat.cbcb.umd.edu/; v2.0.9) with default settings. The abundance of RNA-seq reads aligned (at least three reads) to an eTM sequence indicates the expression level of the eTM. To investigate the conservation of the soybean eTMs in other oil crops or their progenitor species, including *R. communis*, *L. usitatissimum*, *B. rapa*, and *G. raimondii*, miRNA pairing sites of the *G. max* eTMs were used to Blast against the genomes of the four species (BlastN, 1e-1). Alignment of eTMs from soybean and other species was achieved using sequences including the miRNA pairing sites along with their flanking regions of the predicted eTMs (100 bp in total).

2.4 Identification of putative phasiRNA triggers

PhasiRNA-producing (*PHAS*) loci were identified in the intergenic regions of the *G. max* genome using the following algorithm previously described by Howell et al (Howell et al., 2007).

$$\mathrm{P} = In\left[\left(1 + \sum_{i=1}^{8} k_i\right)^{n-2}\right], \mathrm{P} > 0,$$

(where n = number of phase cycle positions occupied by at least one small RNA read within an eight-cycle window, and k = total number of reads for all small RNAs falling into a given phase within an eight-cycle window). In this study, all phasiRNA-producing

loci were considered as *PHAS* loci, although they were located in the non-coding regions of the *G. max* genome. For a miRNA to be considered as a putative phasiRNA trigger, (1) it has to be able to bind to a *PHAS* locus within the regions generating phased siRNAs or their flanking 200-nt regions; (2) the putative miRNA cleavage site (corresponding to the position between the 10^{th} and 11^{th} nucleotides counting from the 5' end of a miRNA) should be at a position corresponding to a phase register. (Liu et al., 2013) The expression level of a *PHAS* gene was measured by the abundance of siRNA reads (RPM) aligned to all phases of the *PHAS* locus.

2.5 Analysis of potential genes related to lipid biosynthesis in soybean

The lipid biosynthesis related genes in Arabidopsis were downloaded from http://aralip.plantbiology.msu.edu/ and were then searched against the *G. max* genome to find their homologous soybean genes using BlastP with an E-value < 1e-5 and identity > 50% (Li-Beisson et al., 2013;Wang et al., 2014).

2.6 Target prediction and validation by degradomic data

Targets of soybean miRNAs were predicted by search against all annotated transcripts of *G. max* using psRNATarget (http://plantgrn.noble.org/psRNATarget/) with the default settings (maximum expectation: 3.0; length for complementary scoring: 20 bp; target accessibility-allowed maximum energy to unpair the target site: 25.0; flanking length around target site for target accessibility analysis: 17 bp in upstream and 13 bp in downstream; range of central mismatch leading to translational inhibition: 9-11 nt) (Dai and Zhao, 2011). These settings, except for maximum expectation that was set as 1 to reduce the false positive prediction rate, were also used in prediction of targets of phasiRNAs. To make the miRNA target prediction more accurate, psRobot was also used with the score of three (Wu et al., 2012). The same results obtained from these two programs were adopted. The soybean homologs of Arabidopsis genes related to lipid biosynthesis that were predicted to be targets of miRNAs or phasiRNAs were considered as candidates involved in lipid biosynthesis and regulated by miRNAs or phasiRNAs in soybean. Publicly available degradome sequencing data (Table S1) were used to validate the predicted targets. A predicted target was considered as cleaved by a miRNA or phasiRNA if the predicted cleavage site had degradome reads perfectly aligned to, *i.e*. the 5' ends of degradome reads were exactly aligned with the predicted cleavage site of a miRNA or phasiRNA.

3. Results

3.1 Identification of endogenous target mimics

Intergenic non-coding sequences in the soybean genome (Williams 82) were selected to predict putative eTMs for all 554 soybean miRNAs. Of the 554 soybean miRNAs, 334 (60.3%) had eTMs predicted. In total, we predicted 10,410 eTMs for 144 conserved miRNAs and 32,555 for 190 soybean-specific miRNAs (Table S2). As an example, sequence alignment between the soybean-specific gma-miR1522 and one of its eTMs, gma-eTM1522-2, was shown in Figure 1A, which showed a typical pairing pattern

between a miRNA and its eTM, *i.e.* three unmatched nucleotides in the eTM at the region corresponding to the 10th to 12th positions from the 5' end of gma-miR1522 (Figure 1A). A number of eTMs were predicted for gma-miR1522 and 22 of them were expressed in at least one of the tissues or organs examined in this study (Tables S1, S2). Sequence alignment of these 22 expressed eTMs showed that they were conserved only in the region corresponding to the predicted miRNA binding sites (*i.e.* the target mimic sites) (Figure 1B). Sequences containing a well conserved target mimic site for gma-miR1522 were also found in the four additional species (*R. communis*, *L. usitatissimum*, *B. rapa*, and *G. raimondii*) examined in this study. Similarly, only the target mimic sites but not their flanking regions were conserved in these predicted eTMs (Figure 1C), consistent with the observation previously reported (Wu et al., 2013). Prediction of conserved eTMs for gma-miR1522 in all four species examined in this study suggests that miR1522 could be present, although it has not been reported, in these species.

Expression of the predicted eTMs is a prerequisite for them to be functional. To examine whether the predicted eTMs are expressed in soybean, we analyzed 20 published RNA-Seq datasets generated from a wide range of organs or tissues, including vegetative and reproductive tissues as well as developing seeds (Table S1). In total, 457 eTMs for 126 miRNAs (144 eTMs for 61 conserved miRNAs and 313 for 65 soybean-specific miRNAs) were found to be expressed in at least one of the examined tissues (Table 1, Table S2). These 457 eTMs for miRNAs with high confidence were used for our further investigation. Of these 457 expressed eTMs, 285 had their homologs found in at least one of the other four species examined in this study. For example, the homolog of eTM1522 could be found in the all four species (Figure 1C). To rule out the possibility that the expressed eTMs are potential miRNA targets, we analyzed seven published degradome datasets and found that no degradome read was assigned to the miRNA pairing regions of the 457 eTMs, demonstrating no miRNA-mediated cleavage in these eTMs and suggesting that they could function as mRNA decoys. To know the potential functionality of these eTMs, we performed Gene Ontology (GO) analysis for the predicted targets of the 126 miRNAs with expressed eTMs. As expected these miRNA targets were found to be involved in a wide range of biological progresses (Table S3); we thus expect that the expressed eTMs could also play a role in the related biological progresses. Taken together, the above results suggest that the functions of a number of miRNAs were potentially regulated by eTMs in soybean.

3.2 eTMs for miRNAs targeting genes related to lipid metabolism

To investigate a role of miRNAs and their eTMs in lipid metabolism in soybean, we first identified soybean homologs of Arabidopsis genes involved in lipid metabolism. This analysis found 1,507 genes with a potential role in lipid metabolism in soybean. We then predicted potential targets of all soybean miRNAs (in total 554, including 281 conserved and 273 soybean-specific miRNAs) downloaded from miRBase (Release 20). Based on these two analyses, we found that 89 soybean homologs of Arabidopsis genes involved in lipid biosynthesis were predicted targets of 97 miRNAs (45 conserved and 52 soybean-specific miRNAs) (Table 1). Using the publicly available degradome data we found that 25 of these genes were cleaved by miRNAs from 18 miRNA families (Table S4, Figure S1). According to gene annotation, these genes are involved in a diverse

pathways related to lipid metabolism, such as fatty acid synthesis, elongation and degradation as well as lipid trafficking.

Of the 457 expressed eTMs, 119 were eTMs for 33 miRNAs predicted to target genes involved in lipid metabolism, including 37 eTMs for 14 conserved miRNAs and 82 eTMs for 19 soybean-specific miRNAs (Tables S5, S6). Some of these eTMs were expressed in most of the tissues or organs examined whereas the majority of these eTMs were specifically expressed in certain tissue(s). For example, gma-eTM1522-17 seemed to be universally expressed while expression of 33, 16 and 10 eTMs was only found in cotyledon, seed coat and endosperm at the early stage of seed maturation, respectively (Figure 2). Some miRNAs had multiple expressed eTMs that showed differential expression pattern in different tissues. For instance, we identified 22 expressed eTMs for gma-miR1522, which was predicted to target a gene encoding ABC transporter (Glyma03g29150, homolog of At3g21090 that is required for lipid transport; (Pighin et al., 2004); Table S5). Four of the 22 eTMs were expressed in four or more different samples shown in Figure 2, whereas 15 of the 22 eTMs were only expressed in one of the samples (Figure 2). Differential expression of different eTM members for the same miRNA was also observed for eTMs of other miRNAs, such as gma-miR1530 (10 eTMs) and gma-miR504 (9 eTMs), which were predicted to target a phosphatidylinositol kinase gene (Glyma04g39150, homolog of At5g64070 that is involved in phospholipid signaling; (Li-Beisson et al., 2013) and genes (Glyma04g06630 and Glyma06g06720) related to triacylglycerol biosynthesis (Table S5), respectively. Additionally, 65 of the 119 expressed eTMs for miRNAs predicted to target genes involved in lipid metabolism had a homolog in at least one of the other four species used in this study (Table S7; eTMs for miR1522 as an example was shown in Figure 1C). The conservation of eTMs further supported the potential functionality of these eTMs in lipid metabolism.

3.3 Identification of miRNA-mediated phasiRNA-producing loci

PhasiRNA-producing (*PHAS*) loci of the *G. max* (Williams 82) genome were identified based on small RNA reads from eight small RNA libraries (Table S1) and the algorithm described by Howell et al (Howell et al., 2007). A large number of *PHAS* loci generating 21-nt or 24-nt phasiRNAs were detected in the *G. max* genome (phase score > 1.4). Even when using a stringent phase score (phase score > 20), 1,573 of 21-nt and 4,681 of 24-nt *PHAS* loci could still be identified (Table S8). Of these *PHAS* loci, 984 (262 of 21-nt and 722 of 24-nt *PHAS* loci) had at least a miRNA binding site predicted within the regions generating phased siRNAs and/or their flanking 200-nt regions (Table S9). Some of the identified phasiRNA triggers have been previously identified in other studies, such as miR390, miR156, miR2118, miR393, miR1508, miR1510, and miR1514 (Zhai et al., 2011;Hu et al., 2013). Of the 984 *PHAS* loci, 157 and 827 were triggered by conserved and soybean-specific miRNAs, respectively, and phasiRNAs from 49 such *PHAS* loci were predicted to target genes related to lipid metabolism (Tables 1, S5, S9). For example, gma-miR1520j was predicted to trigger production of 24-nt phasiRNAs from a locus (*PHAS1520j-3*) located in soybean chromosome 17, and a phasiRNA from this locus was predicted to target Glyma03g31570, which encodes an acylhydrolase involved in oxylipin metabolism (Figure 3). Of these 49 *PHAS* genes, 12 were cleaved by their miRNA triggers according to degradome data (Table S9).

The expression levels of the 49 *PHAS* genes producing phasiRNAs targeting genes related to lipid metabolism were investigated based on published small RNA data (Figure 4, Tables S1, S10). Of the 49 *PHAS* genes, nine seemed to be expressed in two or more of the eight tissue samples shown in Figure 4, although the expression levels were low in most samples. One exception was the *PHAS4388-5* gene, which was targeted by gma-miR4388 to produce phasiRNA targeting Glyma12g28470 that encodes an acyltransferase. This *PHAS* gene was highly expressed in seed but lowly expressed in leaf in cultivar Heinong44 (Figure 4). Interestingly, none of the 49 *PHAS* genes seemed to be expressed in stems of soybean although some of these *PHAS* genes were relatively highly expressed in roots and/or leaves.

4. Discussion

4.1 A large number of eTMs and *PHAS* genes in soybean

It has become clear that numerous non-coding RNA transcripts interact with miRNAs and are part of the network regulating development and stress responses in both plants and animals (Tay et al., 2014). Because of their ability to sequester miRNAs away from their cleavable targets, RNA molecules with miRNA binding sites but un-cleavable by miRNAs have been reported with different terminology, such as 'miRNA sponges/decoys', 'endogenous target mimics' or 'competing endogenous RNA' (Banks et al., 2012;Wu et al., 2013;Gupta, 2014). In this study, we found that the functionalities of a large number of miRNAs could be potentially regulated by eTMs in soybean. Our results together with previously published results further suggest that eTMs could be widespread regulators of miRNA functions in plants (Wu et al., 2013). Among the 42,965 computationally predicted eTMs in the soybean genome, only 457 eTMs were found to be expressed. It was partly due to that the RNA-Seq data we used only included samples from different tissues and developmental stages of soybean. The functions of eTMs are related to their corresponding miRNAs, many of which function in stress responses. Therefore，RNA-Seq data from samples with stress treatment should provide expression evidence for more eTMs. However, the huge discrepancy between computationally-predicted eTMs and those with expression support also demonstrated the limitation of the computational prediction for eTMs. The expression evidence is thus important for the identification of authentic eTMs. Another kind of interaction between miRNAs and non-cding RNAs involves generating phasiRNAs from the miRNA-targeted non-coding RNAs. The well known phasiRNAs generated from miRNA-targeted non-coding RNAs are tasiRNAs. *TAS* gene has been demonstrated to play a role in diverse biological processes, such as development by targets auxin response factor (ARF) and MYB protein family and disease resistance by nucleotide binding leucine-rich repeat NB-LRR (Williams et al., 2005;Zhai et al., 2011). In addition to *TAS* genes, we showed here that numerous intergenic regions in the soybean genome are targeted by miRNAs to produce phasiRNAs, which are able to further target other mRNAs. This result suggested that miRNA-mediated production of phasiRNAs may have a much more profound effect on biological processes in soybean than we previously thought.

More eTMs (32,555 versus 10,410 eTMs based on computational prediction; 313 versus 144 eTMs with expression support) were found for the 190 soybean-specific miRNAs than for the 144 conserved miRNAs. Similarly, we also found that soybean *PHAS* genes are more likely to be targeted by soybean-specific miRNAs than by conserved miRNAs (827 versus 157 *PHAS* genes). This phenomenon may suggest that eTMs and phasiRNAs are two types of important regulators co-evolved with soybean-specific miRNAs to regulate soybean-specific biological processes, and that the co-evolution of the non-coding RNA network in soybean may be a result of species-specific adaptations. In addition, we found that some biological processes were significantly enriched in the predicted target genes of miRNAs that have eTMs with expression evidence, suggesting a specified role of the interaction between miRNAs and their eTMs in these biological processes (Table S3). For example, the genes with GO categories related to biological regulation were significantly enriched, indicating the role of eTMs in gene regulation through the network involved miRNAs. Some genes related to stress response and cell death were significantly enriched, suggesting that eTMs might play key roles in these two biological processes through interaction with miRNAs.

4.2 The non-coding RNA network involved in lipid metabolism in soybean

At least 120 enzymatic reactions and more than 600 genes are involved in the 24 pathways related to Acyl-lipid metabolism in Arabidopsis (http://aralip.plantbiology.msu.edu/) (Li-Beisson et al., 2013). Although the pathways and protein coding genes associated with lipid biosynthesis in higher plants have been largely uncovered, the roles of the non-coding RNAs in lipid biosynthesis are still poorly understood. In this study, we investigated the potential roles of the non-coding RNA network, including miRNAs and their partners, *i.e.* eTMs and miRNA-triggered *PHAS* genes, in lipid metabolism in soybean. Understanding the functions of the miRNA network in regulating lipid metabolism in soybean will be of great value for the cultivation of soybean cultivars with increased oil content. Overall, our results revealed that lipid metabolism related genes in soybean are potentially directly regulated by 28 miRNAs with degradome data support, and that nine of these 28 miRNAs are potentially further regulated by a number of eTMs with 40 of them supported by expression data. The lipid metabolism pathways regulated by the network involving these nine miRNAs and 40 eTMs includes fatty acid synthesis, elongation, degradation, oxylipin metabolism, and phospholipid signaling (Table S5). In addition, 33 miRNAs were found to trigger production of phasiRNAs from 49 *PHAS* genes, which were able to target lipid biosynthesis related genes (Figure 5). We found that the lipid metabolism related genes potentially regulated by miRNAs and their partners in soybean were Arabidopsis homologs involved in 23 of the 24 pathways related to Acyl-lipid metabolism (Table S5). These results provided bioinformatical evidences for the hypothesis that lipid metabolism in soybean is regulated by a complex non-coding RNA network including miRNAs, eTMs and phasiRNAs.

This hypothesis was supported by several pieces of evidence. Firstly, of the 97 miRNAs predicted to target lipid biosynthesis related genes, 28 were confirmed to cleave their targets based on the data from seven publicly available degradome libraries (Table S4). Secondly, 119 eTMs for miRNAs predicted to target genes related to lipid metabolism

were found to be expressed and most of them had a tissue-specific expression pattern (Figure 2); 55% of them are evolutionarily conserved in their target mimic sites in the other four species examined in this study. Thirdly, majority of the 49 *PHAS* genes generating phasiRNAs that were predicted to target lipid biosynthesis related genes also showed a tissue-specific expression pattern (Figure 4), and miRNA-mediated cleavage evidence was found for 12 of the 49 *PHAS* genes based on the publicly available degradome data (Table S9). Fourthly, some of the genes related to lipid metabolism might be regulated by only one component of the network but some could be controlled by a cascade of the network or even all three types of non-coding RNAs, *i.e.* miRNA, eTM and phasiRNA. Of the 18 miRNA families targeting lipid biosynthesis genes that were validated by degradome data, at least three (gma-miR1520j, gma-miR4388 and gma-miR4992) were also able to target *PHAS* genes to produce phasiRNAs that in turn to target genes involved in lipid metabolism (Tables S4, S9). Furthermore, at least nine (gma-miR1508b, gma-miR1520p, gma-miR1530, gma-miR2108b, gma-miR394a-3p, gma-miR395b, gma-miR396b-5p, gma-miR5041 and gma-miR5769) of the 18 miRNA families had expressed eTMs (Table S2), suggesting that the functionality of these miRNAs could be attenuated in certain tissues.

In this study, our focus was on non-coding RNAs interacting with miRNAs in soybean; however, in addition to non-coding RNAs, our analyses also found that 120 protein-coding genes were potentially targeted by miRNAs to generate phasiRNAs (Table S9). Additionally, we only used intergenic regions in eTM prediction in this study. In fact, protein-coding, intronic, and antisense sequences all could function as eTMs (Ponting et al., 2009;Wu et al., 2013). Furthermore, it is our further interest to know whether plants contain other kinds of RNA molecules, such as circular RNAs and pseudogene competing endogenous RNAs reported in humans and animals (Tay et al., 2014), and whether they interact with miRNAs to regulate biological processes in plants.

5. Author contributions

LF conceived and designed the study. CY, HX, ES, YL, YW, SY, and JQ analyzed the data. CY, QZ, and LF wrote the paper.

6. Acknowledgments

This work was supported by the National Basic Research Program of China (2011CB109306), Ministry of Science and Technology of China (No. 2013IM030700), and the National Natural Science Foundation of China (31301046). We also thank Jiangsu Collaborative Innovation Center for Modern Crop Production (JCIC-MCP).

Reference

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. doi: 10.1016/j.cell.2005.04.004.

Axtell, M.J., Jan, C., Rajagopalan, R., and Bartel, D.P. (2006). A two-hit trigger for siRNA biogenesis in plants. *Cell* 127**,** 565-577. doi: 10.1016/j.cell.2006.09.032.

Banks, I.R., Zhang, Y., Wiggins, B.E., Heck, G.R., and Ivashuta, S. (2012). RNA decoys: an emerging component of plant regulatory networks? *Plant Signal Behav* 7**,** 1188-1193. doi: 10.4161/psb.21299.

Collakova, E., Aghamirzaie, D., Fang, Y., Klumas, C., Tabataba, F., Kakumanu, A., Myers, E., Heath, L.S., and Grene, R. (2013). Metabolic and transcriptional reprogramming in developing soybean (Glycine max) embryos. *Metabolites* 3**,** 347-372. doi: 10.3390/metabo3020347.

Dai, X., and Zhao, P.X. (2011). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res* 39**,** W155-159. doi: 10.1093/nar/gkr319.

Fei, Q., Xia, R., and Meyers, B.C. (2013). Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *Plant Cell* 25**,** 2400-2415. doi: 10.1105/tpc.113.114652.

Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J.A., and Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat Genet* 39**,** 1033-1037. doi: 10.1038/ng2079.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40**,** D1178-1186. doi: 10.1093/nar/gkr944.

Gupta, P. (2014). Competing endogenous RNA (ceRNA): a new class of RNA working as miRNA sponges. *Curr Sci India* 106**,** 823-830.

Howell, M.D., Fahlgren, N., Chapman, E.J., Cumbie, J.S., Sullivan, C.M., Givan, S.A., Kasschau, K.D., and Carrington, J.C. (2007). Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting. *Plant Cell* 19**,** 926-942. doi: 10.1105/tpc.107.050062.

Hu, Z., Jiang, Q., Ni, Z., Chen, R., Xu, S., and Zhang, H. (2013). Analyses of a *Glycine max* degradome library identify microRNA targets and microRNAs that trigger secondary siRNA biogenesis. *J Integr Plant Biol* 55**,** 160-176. doi: 10.1111/jipb.12002.

Ivashuta, S., Banks, I.R., Wiggins, B.E., Zhang, Y.J., Ziegler, T.E., Roberts, J.K., and Heck, G.R. (2011). Regulation of gene expression in plants through miRNA inactivation. *Plos One* 6. doi: DOI 10.1371/journal.pone.0021330.

Jones-Rhoades, M.W., Bartel, D.P., and Bartel, B. (2006). MicroRNAs and their regulatory roles in plants. *Annu Rev Plant Biol* 57**,** 19-53. doi: 10.1146/annurev.arplant.57.032905.105218.

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. doi: 10.1186/1471-2105-11-S1-S14.

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. doi: 10.1016/j.bbagrm.2011.05.001.

Kulcheski, F.R., De Oliveira, L.F., Molina, L.G., Almerao, M.P., Rodrigues, F.A., Marcolino, J., Barbosa, J.F., Stolf-Moreira, R., Nepomuceno, A.L., Marcelino-Guimaraes, F.C., Abdelnoor, R.V., Nascimento, L.C., Carazzolle, M.F., Pereira, G.A., and Margis, R. (2011). Identification of novel soybean microRNAs involved in abiotic and biotic stresses. *BMC Genomics* 12**,** 307. doi: 10.1186/1471-2164-12-307.

Lardizabal, K., Effertz, R., Levering, C., Mai, J., Pedroso, M.C., Jury, T., Aasen, E., Gruys, K., and Bennett, K. (2008). Expression of *Umbelopsis ramanniana* DGAT2A in seed increases oil in soybean. *Plant Physiol* 148**,** 89-96. doi: 10.1104/pp.108.123042.

Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M., Arondel, V., Bates, P., Baud, S., Bird, D., Debono, A., and Durrett, T. (2013). Acyl-lipid metabolism. *In the Arabidopsis Book* 8**,** e0161. doi: 10.1199/tab.0161.

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. doi: 10.1104/pp.110.156950.

476 Li, H., Dong, Y., Yin, H., Wang, N., Yang, J., Liu, X., Wang, Y., Wu, J., and Li, X. (2011). Characterization
477 of the stress associated microRNAs in *Glycine max* by deep sequencing. *BMC Plant Biol* 11**,** 170.
478 doi: 10.1186/1471-2229-11-170.
479 Libault, M., Farmer, A., Joshi, T., Takahashi, K., Langley, R.J., Franklin, L.D., He, J., Xu, D., May, G., and
480 Stacey, G. (2010). An integrated transcriptome atlas of the crop model Glycine max, and its use in
481 comparative analyses in plants. *Plant J* 63**,** 86-99. doi: 10.1111/j.1365-313X.2010.04222.x.
482 Liu, Y., Wang, Y., Zhu, Q.H., and Fan, L. (2013). Identification of phasiRNAs in wild rice (*Oryza*
483 *rufipogon*). *Plant Signal Behav* 8**,** e25079. doi: 10.4161/psb.25079.
484 Mallory, A.C., and Vaucheret, H. (2006). Functions of microRNAs and related small RNAs in plants. *Nat*
485 *Genet* 38 Suppl**,** S31-36. doi: 10.1038/ng1791.
486 Meyers, B.C., Axtell, M.J., Bartel, B., Bartel, D.P., Baulcombe, D., Bowman, J.L., Cao, X., Carrington, J.C.,
487 Chen, X., Green, P.J., Griffiths-Jones, S., Jacobsen, S.E., Mallory, A.C., Martienssen, R.A.,
488 Poethig, R.S., Qi, Y., Vaucheret, H., Voinnet, O., Watanabe, Y., Weigel, D., and Zhu, J.K. (2008).
489 Criteria for annotation of plant MicroRNAs. *Plant Cell* 20, 3186-3190. doi:
490 10.1105/tpc.108.064311.
491 Pighin, J.A., Zheng, H., Balakshin, L.J., Goodman, I.P., Western, T.L., Jetter, R., Kunst, L., and Samuels,
492 A.L. (2004). Plant cuticular lipid export requires an ABC transporter. *Science* 306**,** 702-704. doi:
493 10.1126/science.1102331.
494 Ponting, C.P., Oliver, P.L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell*
495 136**,** 629-641. doi: 10.1016/j.cell.2009.02.006.
496 Rajagopalan, R., Vaucheret, H., Trejo, J., and Bartel, D.P. (2006). A diverse and evolutionarily fluid set of
497 microRNAs in *Arabidopsis thaliana*. *Genes Dev* 20**,** 3407-3425. doi: 10.1101/gad.1476406.
498 Rubio-Somoza, I., and Weigel, D. (2011). MicroRNA networks and developmental plasticity in plants.
499 *Trends Plant Sci* 16**,** 258-264. doi: 10.1016/j.tplants.2011.03.001.
500 Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J.,
501 Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K.,
502 Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S., Goodstein, D., Barry, K.,
503 Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N., Joshi, T., Libault, M.,
504 Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J.,
505 Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C., and Jackson, S.A. (2010). Genome
506 sequence of the palaeopolyploid soybean. *Nature* 463**,** 178-183. doi: 10.1038/nature08670.
507 Shamimuzzaman, M., and Vodkin, L. (2012). Identification of soybean seed developmental stage-specific
508 and tissue-specific miRNA targets by degradome sequencing. *BMC Genomics* 13**,** 310. doi:
509 10.1186/1471-2164-13-310.
510 Song, Q.X., Liu, Y.F., Hu, X.Y., Zhang, W.K., Ma, B.A., Chen, S.Y., and Zhang, J.S. (2011). Identification
511 of miRNAs and their target genes in developing soybean seeds by deep sequencing. *Bmc Plant*
512 *Biol* 11. doi: 10.1186/1471-2229-11-5.
513 Song, Q.X., Lu, X., Li, Q.T., Chen, H., Hu, X.Y., Ma, B., Zhang, W.K., Chen, S.Y., and Zhang, J.S. (2013).
514 Genome-wide analysis of DNA methylation in soybean. *Mol Plant* 6**,** 1961-1974. doi:
515 10.1093/mp/sst123.
516 Subramanian, S., Fu, Y., Sunkar, R., Barbazuk, W.B., Zhu, J.K., and Yu, O. (2008). Novel and
517 nodulation-regulated microRNAs in soybean roots. *BMC Genomics* 9**,** 160. doi:
518 10.1186/1471-2164-9-160.
519 Tay, Y., Rinn, J., and Pandolfi, P.P. (2014). The multilayered complexity of ceRNA crosstalk and
520 competition. *Nature* 505**,** 344-352. doi: 10.1038/nature12986.
521 Todesco, M., Rubio-Somoza, I., Paz-Ares, J., and Weigel, D. (2010). A collection of target mimics for
522 comprehensive analysis of microRNA function in *Arabidopsis thaliana*. *PLoS Genet* 6**,** e1001031.
523 doi: 10.1371/journal.pgen.1001031.
524 Turner, M., Yu, O., and Subramanian, S. (2012). Genome organization and characteristics of soybean
525 microRNAs. *BMC Genomics* 13**,** 169. doi: 10.1186/1471-2164-13-169.
526 Wang, L., Yu, S., Tong, C., Zhao, Y., Liu, Y., Song, C., Zhang, Y., Zhang, X., Wang, Y., Hua, W., Li, D., Li,
527 F., Yu, J., Xu, C., Han, X., Huang, S., Tai, S., Wang, J., Xu, X., Li, Y., Liu, S., and Varshney, R.K.
528 (2014). Genome sequencing of the high oil crop sesame provides insight into oil biosynthesis.
529 *Genome Biol* 15**,** R39. doi: 10.1186/gb-2014-15-2-r39.
530 Wang, Y., Li, P., Cao, X., Wang, X., Zhang, A., and Li, X. (2009). Identification and expression analysis of
531 miRNAs from nitrogen-fixing soybean nodules. *Biochem Biophys Res Commun* 378**,** 799-803. doi:

10.1016/j.bbrc.2008.11.140.
Williams, L., Carles, C.C., Osmont, K.S., and Fletcher, J.C. (2005). A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the Arabidopsis ARF2, ARF3, and ARF4 genes. *Proc Natl Acad Sci U S A* 102**,** 9703-9708. doi: 10.1073/pnas.0504029102.
Wong, C.E., Zhao, Y.T., Wang, X.J., Croft, L., Wang, Z.H., Haerizadeh, F., Mattick, J.S., Singh, M.B., Carroll, B.J., and Bhalla, P.L. (2011). MicroRNAs in the shoot apical meristem of soybean. *J Exp Bot* 62**,** 2495-2506. doi: 10.1093/jxb/erq437.
Wu, H.J., Ma, Y.K., Chen, T., Wang, M., and Wang, X.J. (2012). PsRobot: a web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res* 40, W22-28. doi: 10.1093/nar/gks554.
Wu, H.J., Wang, Z.M., Wang, M., and Wang, X.J. (2013). Widespread long noncoding RNAs as endogenous target mimics for microRNAs in plants. *Plant Physiol* 161**,** 1875-1884. doi: 10.1104/pp.113.215962.
Xiao, B., Yang, X., Ye, C.Y., Liu, Y., Yan, C., Wang, Y., Lu, X., Li, Y., and Fan, L. (2014). A diverse set of miRNAs responsive to begomovirus-associated betasatellite in *Nicotiana benthamiana*. *BMC Plant Biol* 14**,** 60. doi: 10.1186/1471-2229-14-60.
Yan, J., Gu, Y., Jia, X., Kang, W., Pan, S., Tang, X., Chen, X., and Tang, G. (2012). Effective small RNA destruction by the expression of a short tandem target mimic in Arabidopsis. *Plant Cell* 24**,** 415-427. doi: 10.1105/tpc.111.094144.
Yoshikawa, M., Peragine, A., Park, M.Y., and Poethig, R.S. (2005). A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes Dev* 19**,** 2164-2175. doi: 10.1101/gad.1352605.
Zeng, H.Q., Zhu, Y.Y., Huang, S.Q., and Yang, Z.M. (2010). Analysis of phosphorus-deficient responsive miRNAs and cis-elements from soybean (*Glycine max* L.). *J Plant Physiol* 167**,** 1289-1297. doi: 10.1016/j.jplph.2010.04.017.
Zeng, Q.Y., Yang, C.Y., Ma, Q.B., Li, X.P., Dong, W.W., and Nian, H. (2012). Identification of wild soybean miRNAs and their target genes responsive to aluminum stress. *Bmc Plant Biol* 12. doi: 10.1186/1471-2229-12-182.
Zhai, J.X., Jeong, D.H., De Paoli, E., Park, S., Rosen, B.D., Li, Y.P., 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. doi: 10.1101/gad.177527.111.
Zhao, Y.T., Wang, M., Fu, S.X., Yang, W.C., Qi, C.K., and Wang, X.J. (2012). Small RNA profiling in two Brassica napus cultivars identifies microRNAs with oil production- and development-correlated expression and new small RNA classes. *Plant Physiol* 158**,** 813-823. doi: 10.1104/pp.111.187666.
Zhu, H., Xia, R., Zhao, B., An, Y.Q., Dardick, C.D., Callahan, A.M., and Liu, Z. (2012). Unique expression, processing regulation, and regulatory network of peach (*Prunus persica*) miRNAs. *BMC Plant Biol* 12**,** 149. doi: 10.1186/1471-2229-12-149.

572 **Table 1**. Number of miRNAs and their partners (eTMs and *PHAS* loci) in soybean

Type of miRNAs	Number of miRNAs (lipid[a])	Number of expressed eTMs for miRNAs (lipid[b])	Number of *PHAS* loci with a miRNA trigger identified (lipid[c])
Conserved	281 (45)	144 (37)	157 (3)
Soybean-specific	273 (52)	313 (82)	827 (46)
Total	554 (97)	457 (119)	984 (49)

573 [a]Number of miRNAs that were predicted to target lipid biosynthesis related genes.
574 [b]Number of eTMs for miRNAs that were predicted to target lipid biosynthesis related
575 genes.
576 [c]Number of *PHAS* loci producing phasiRNAs that were predicted to target lipid
577 biosynthesis related genes.
578
579

Figure legends

Figure 1. Endogenous target mimics (eTMs) for gma-miR1522 in soybean. (A) The predicted base-pairing pattern between gma-miR1522 and one of its eTMs; (B) Sequence alignment of the 22 expressed eTMs for gma-miR1522; (C) Sequence alignement of eTMs for miR1522 in *G. max* and other four species (lus: *L. usitatissimum*; rco: *R. communis*; bra: *B. rapa*; and gra: *G. raimondii*).

Figure 2. Expression patterns of the eTMs for miRNAs with a potential role in lipid biosynthesis. Expression values were $\log_2$-transformed.

Figure 3. A *PHAS* locus (Gm17:18214876…18215211) targeted by gma-miR1520j. (A) The siRNA abundance and phasing score of the phased siRNAs generated from the *PHAS* locus *PHAS1520j-3*; (B) Sequence segment of *PHAS1520j-3*. The gma-miR1520j binding site is aligned with gma-miR1520j (predicted target location: Gm17:18214862…18214885) and phases are separated by asterisks; (C) Sequence alignment between a phasiRNA (Gm17: 18215043…18215066) derived from *PHAS1520j-3* locus and its predicted target (Glyma03g31570, a gene encoding acylhydrolase involved in oxylipin metabolism).

Figure 4 Expression patterns of the *PHAS* genes producing phasiRNAs that were predicted to target lipid biosynthesis related genes in soybean.

Figure 5. A model for the non-coding RNA network involved in lipid biosynthesis in soybean. The network contains 28 miRNAs directly targeting lipid biosynthesis related genes with degradome data support and 33 miRNAs predicted to target 49 *PHAS* loci that produced phasiRNAs to target 12 lipid biosynthesis related genes. Of the 28 miRNAs directly targeting lipid biosynthesis related genes, nine have predicted eTMs supported by expression data.

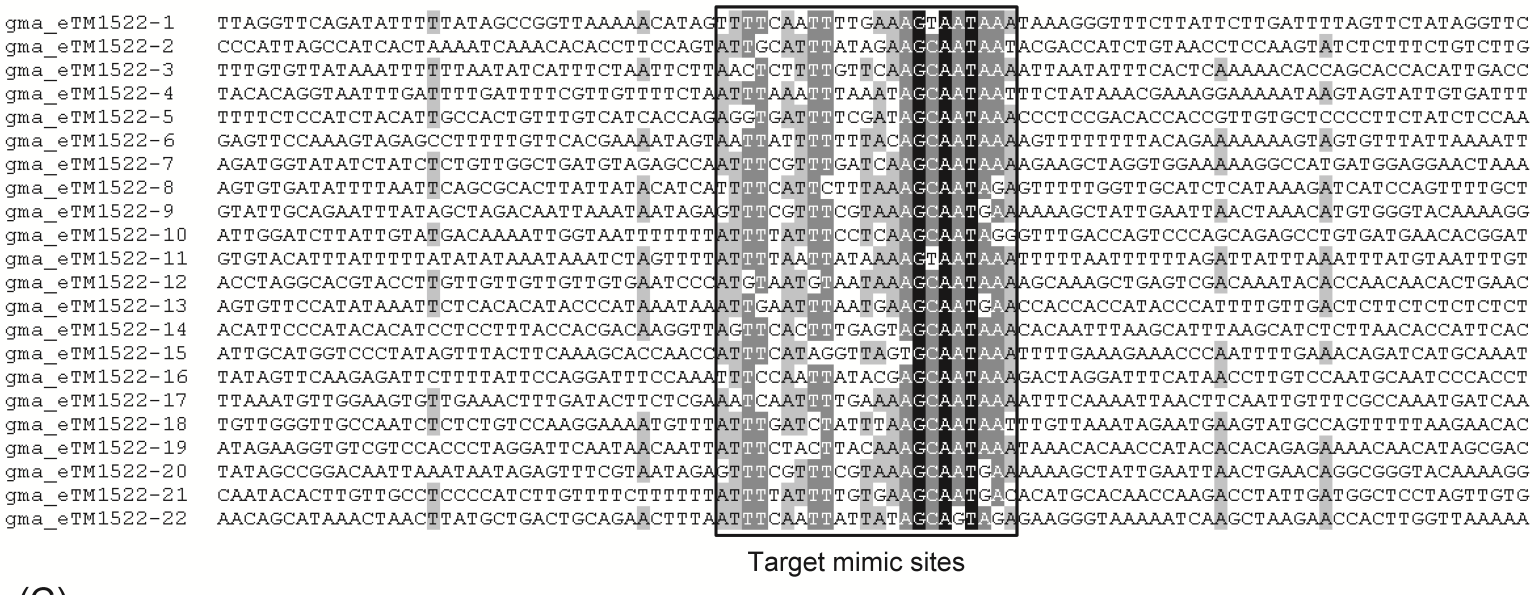
Figure 1.TIF

Fig. 1

(A)

```
gma-miR1522     5′ UUUAUUGCUU--A-AAAUGAAAU 3′
                   o|||||||||  | |||||||||
gma-eTM1522-2   3′ TAATAACGAAGATATTTACGTTA 5′
```

(B)



(C)

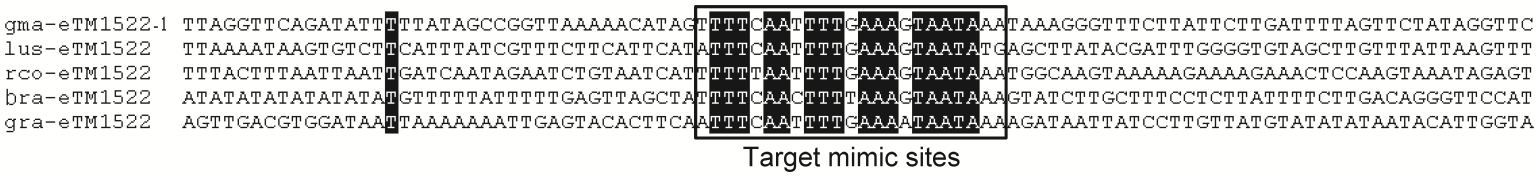


Fig. 2

Figure 2.TIF

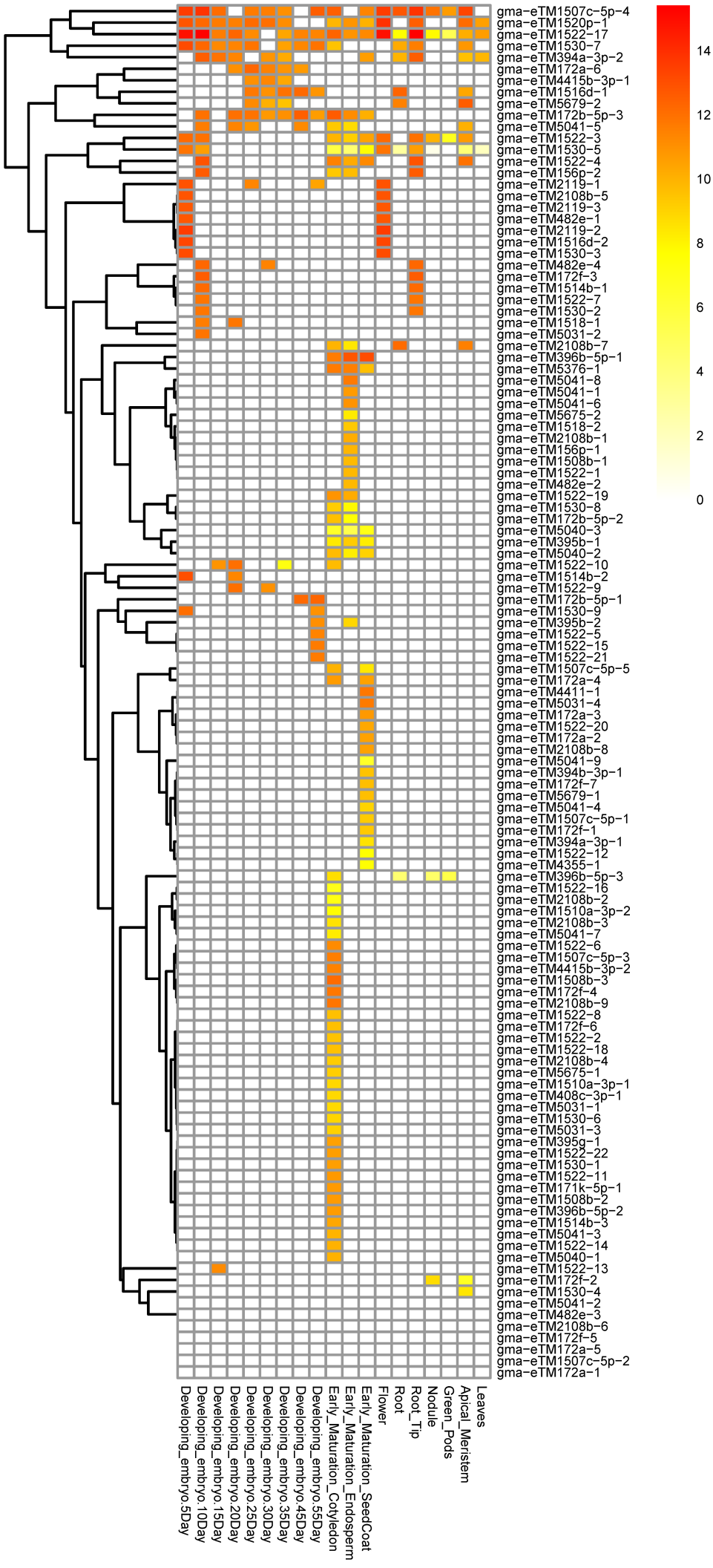
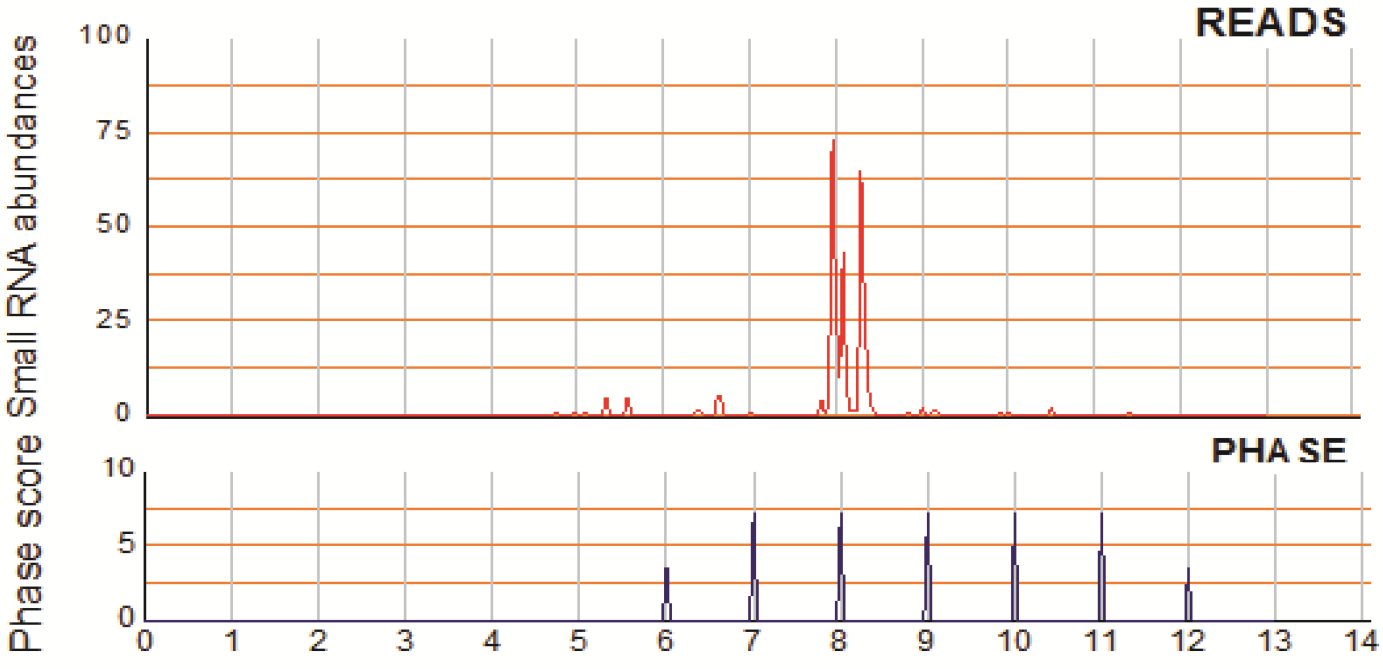


Fig. 3

(A)



(B)

```
gma-miR1520j      3′AACUAACAGUACACAGUGCAAGAA 5′
                    |||| |||||| ||||||||||||
PHAS              5′TTGACTGTCATATGTCACGTTCTTATTGGATGACGA
                  CTGTCACGTGTCATGTTCTGATTGGATGTCAACTCCAT
                  GAAAGATGAAATGCATTGCTGTTTTCATTGACCAATTA
                  GAATATGACATGTGACAGTCATCATCCAATCAGAACGT
                  GATACGTGACAGTTAACATCACTTGTTAACGGTTAATG
                  TCTAATTGTGGCCTGGAGTACCAACATTGCATGATTTT
                  ACAA 3′
```

(C)

```
phasiRNA          5′UGUCAACUCCAUGAAAGAUGAAAU 3′
                    |||||||||||||||||||||||||
Glyma03g31570     3′ACAGTTGAGGTACTTTCTACTTTA 5′
```

Fig. 4

Figure 4.TIF

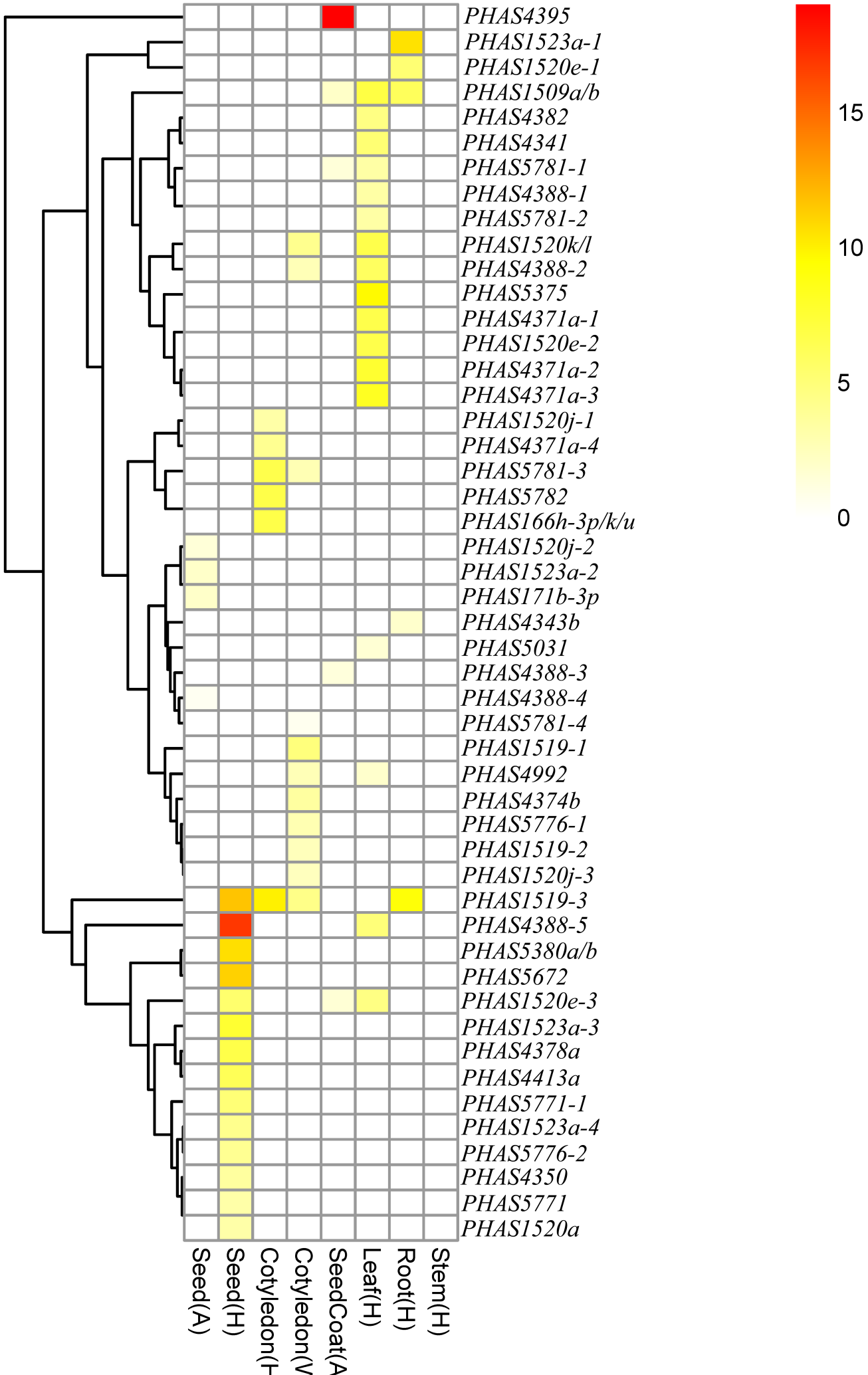


Figure 5.TIF

Fig. 5