

IMPROVE SMALL RNA-MEDIATED GENE SILENCING IN SOYBEAN BY USING
GmFAD3 AS A TEST MODEL

A Dissertation

Presented to

The Faculty of the Graduate School

At the University of Missouri

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

SHA LU

Dr. Zhanyuan Zhang, Dissertation Supervisor

JULY 2014

The undersigned, appointed by the dean of the Graduate School,
have examined the Dissertation entitled

IMPROVE SMALL RNA-MEDIATED GENE SILENCING IN SOYBEAN BY USING
GmFAD3 AS A TEST MODEL

Presented by Sha Lu

A candidate for the degree of

Doctor of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

Dr. Zhanyuan Zhang

Dr. Kristin Bilyeu

Dr. Jim Schoelz

Dr. Thomas Quinn

ACKNOWLEDGEMENTS

I would like to first thank my dissertation supervisor, Dr. Zhanyuan Zhang for giving me the opportunity to join his research group, and for his continuous advising and support during the entire course of this research. I would also like to thank Dr. Kristin Bilyeu and her lab members for their collaboration and contribution to my dissertation projects. My sincere thanks also goes to Dr. Jim Schoelz and Dr. Thomas Quinn for serving on my doctoral committee. I thank Dr. William Spollen and Ning Zhang to provide the Bioinformatics support of this project and my labmates Liwen Zhou, Sandra Valdes, Muruganantham Mookkan and Neng Wan for helping me with soybean transformation and taking care of transgenic plants in the Greenhouse. I would like to give my very special thanks to Jikai Liu, for her support, and most importantly, her friendship during my graduate studies at University of Missouri.

TABLE OF CONTENTS

Acknowledgements.....	ii
List of Figures.....	ix
List of Tables.....	xi
CHAPTER 1.....	1
Literature Review.....	1
THE DISCOVERY OF RNA SILENCING.....	2
RNA SILENCING PATHWAY IN PLANTS.....	4
The exo-siRNA pathway.....	5
The endo-siRNA pathway.....	6
hc-siRNA.....	6
Secondary siRNA.....	8
NAT-siRNA.....	11
The miRNA pathway.....	12
RNAi DELIVERY IN PLANTS.....	15
Transgene-induced RNA silencing.....	15
Virus-induced gene silencing (VIGS).....	17

APPLICATION OF RNAi IN SOYBEAN.....	18
Metabolic engineering.....	18
Virus resistance.....	20
Gene function analysis via VIGS.....	20
IMPROVEMENT OF SYSTEMS FOR RNAi INDUCTION IN SOYBEA...22	
GmFAD3.....	22
FIGURES.....	25
REFERENCE.....	29
CHAPTER 2.....	45
Knocking-down <i>Glycine max</i> omega 3 fatty acid desaturase (<i>FAD3</i>) gene: A case study revealing siRNA-mediated gene silencing process in crop plants.....	45
ABSTRACT.....	46
INTRODUCTION.....	47
METERIALS AND METHODS.....	51
Plant material and growth conditions.....	51
Fatty acid analysis.....	51
qRT-PCR.....	52
Bisulfite Sequencing.....	53

Small RNA sequencing.....	54
5' RACE.....	55
Statistical analysis.....	56
RESULTS.....	57
Low linolenic acid phenotype of <i>GmFAD3</i> RNAi lines is inheritable...	57
Silencing levels of <i>FAD3A</i> , <i>FAD3B</i> and <i>FAD3C</i> correlate to degrees of sequence homology between inverted repeat and <i>GmFAD3</i> mRNA transcripts in the RNAi lines.....	58
Transgenes are silenced in two of the three RNAi lines.....	60
DNA methylation suppresses transgene expression.....	61
No DNA methylation was detected in endogenous glycinin gen.....	64
Small RNA sequencing results uncovers hpRNA processing patterns in stably transformed RNAi lines.....	65
Overall size distribution of small RNAs.....	65
Size distribution of hpRNA-produced siRNAs.....	66
siRNA distribution along 318-bp IR.....	68
Association of hpRNA-produced siRNAs to differential silencing efficacy of target genes in RNAi lines.....	70

Potential transitivity of Small RNAs.....	71
small RNA encoding part of the transgene.....	73
Small RNA encoding part of the Glycinin and 35S promoter.....	73
Small RNAs share homology with endogenous Glycinin transcript...	75
Small RNAs encoding part of the bar coding sequence.....	76
5' RACE uncovered cleavage events directed by siRNAs.....	77
Putative functional siRNAs.....	79
DISCUSSION.....	81
FIGURES.....	90
TABLES.....	113
LITERATURE CITED.....	126
CHAPTER 3.....	132
Development of atasiRNA-based Vectors for Efficient Gene Silencing in	
Soybean.....	132
ABSTRACT.....	133
INTRODUCTION.....	134
METERIALS AND METHODS.....	138
Plant material and growth conditions.....	138

RNAi vector construction.....	138
cDNA overexpression vector construction.....	139
Plant transformation.....	139
Transgene integration and segregation analysis.....	140
Fatty acid analysis.....	141
Agroinfiltration.....	141
qRT-PCR.....	142
stem–loop RT-PCR.....	143
RESULTS.....	145
Comparison of atasiRNA candidates designed by OligoWalk with putative functional siRNAs experimentally determined by 5’ RACE.....	145
<i>Agrobacterium</i> -mediated transformation did not yield enough transgenic plants.....	146
Fatty acid profile revealed increased linoleic acid content in two of the analyzed RNAi lines.....	147
Test silencing efficacy of atasiRNA constructs using transient expression assay.....	148

miR390 temporal and spacial expression analysis.....	149
DISCUSSION.....	152
FIGURES.....	155
TABLES.....	159
LITERATURE CITED.....	168
VITA.....	174

LIST OF FIGURES

Figure	Page
1.1 Exogenous siRNA pathways in plants.....	26
1.2 Endogenous small RNA pathways in plants.....	27
1.3 Comparison of transgene/virus-induced gene silencing pathways in plants.....	28
2.1 <i>FAD3</i> gene expression level of T ₅ homozygous RNAi lines.....	91
2.2 Alignments of nucleotide sequence of the 318-nt inverted repeat (IR) with partial cDNAs of <i>GmFAD3A</i> , <i>GmFAD3B</i> , and <i>GmFAD3C</i>	92
2.3 hpRNA and <i>bar</i> transcript level of pMUFAD homozygous lines.....	93
2.4 Leaf painting phenotype comparison on seedling of T ₄ S-24-13 and WT.....	94
2.5 DNA methylation analysis of T ₅ homozygous RNAi lines.....	95
2.6 Small RNA size distribution.....	98
2.7 Comparison of 318-bp IR derived siRNAs in RNAi lines.....	99
2.8 318-bp IR-derived siRNAs targeting <i>FAD3B</i> and <i>FAD3C</i>	100
2.9 RNAi-induced transitivity.....	102
2.10 Small RNA encoding part of transgene promoters.....	105
2.11 Small RNAs matching endogenous <i>Glycinin</i> transcript.....	107

2.12	Small RNAs encoding part of the <i>bar</i> coding sequence.....	108
2.13	5'RACE on <i>GmFAD3A</i> and <i>GmFAD3C</i> mRNAs in T ₅ RNAi lines.....	109
2.14	Putative functional siRNAs.....	112
3.1	Schematic presentation of the atasiRNA-producing region of 35S-FAD and GlyP-FAD.....	156
3.2	Relative target gene cDNA levels in agroinfiltrated <i>Nicotiana benthamiana</i> plants.....	157
3.3	Temporal and spacial expression analysis of <i>gma-MIR390</i>	158

LIST OF TABLES

Table	Page
2.1 Sequences of primers used for qRT-PCR analysis.....	114
2.2 Sequences of primers used in Bisulfite sequencing.....	115
2.3 Sequences of gene specific primers used in 5' RACE.....	116
2.4 Fatty acid analysis of T ₃ and T ₅ soybeans for homozygous RNAi lines.....	117
2.5 Summary of distinct small RNA and total small RNA abundance for each RNA template.....	118
2.6 Putative functional siRNAs.....	124
3.1 Sequences of atasiRNA candidates generated from the 318-nt IR.....	160
3.2 Sequences of primers used for the screen of transgenic plants.....	163
3.3 Sequences of primers used for qRT-PCR analysis.....	164
3.4 Sequences of primers used in stem-loop qRT PCR.....	165
3.5 Summary of Agrobacterium-mediated transformation.....	166
3.6 Fatty acid analysis of T ₂ hemizygous seeds from atasiRNAi lines.....	167

CHAPTER 1

Literature Review

THE DISCOVERY OF RNA SILENCING

RNA silencing is one of the gene regulatory mechanisms conserved among almost all eukaryotic organisms, which refers to a collection of RNA-mediated sequence-specific inhibition of gene expression, either at the post-transcriptional or transcriptional level (Frizzi and Huang 2010).

The phenomenon of Post-transcriptional gene silencing (PTGS) phenomenon was first discovered in transgenic petunia (*Petunia hybrida*) plants (Napoli et al. 1990). These plants were modified to overexpress the Chalcone Synthase gene (CHS A) that codes for deep purple flower. But surprisingly, transgenic plants producing white or patchy flowers were obtained instead of dark purple flowers. Somehow both the introduced and endogenous forms of the CHS gene were silenced by the transgene and this phenomenon was then termed as “co-suppression” (Napoli et al. 1990). Similar observations have also been reported for plants engineered to express gene or gene segments derived from the viral genome (Baulcombe 1996). In one such study, transgenic tobacco plants expressing the tobacco etch virus (TEV) coat protein (CP) were initially susceptible to TEV infection, but returned to a "recovered" non-infected state 3-5 weeks later. Molecular analysis of the recovered tissue indicated that the gene silencing occurred at the posttranscriptional level (Lindbo et al. 1993).

The underlying mechanisms responsible for this puzzling observation of homology-dependent gene silencing remained unknown for many years, until Fire and

Mello took the approach of directly testing double-stranded RNA (dsRNA) as the silencing trigger in *Caenorhabditis elegans* and proposed the term RNA interference (RNAi) for the first time (Fire et al. 1998). Later on, the potency of dsRNA to induce gene silencing was also demonstrated in plants (Waterhouse et al. 1998), protozoa (Ngo et al. 1998), and insects (Kennerdell and Carthew 1998).

The second major breakthrough in RNA silencing was the identification and association of small RNA (sRNA) molecules in plants actively undergoing post-transcriptional gene silencing (PTGS). Hamilton and Baulcombe (1999) screened for sRNA species in three types of transgene-induced PTGS and one example of virus-induced PTGS (Hamilton and Baulcombe 1999). sRNA molecules of approximately 25 nucleotides complementary to the targeted mRNA were detected in all four silencing backgrounds. These species of sRNA, typically 21-24 nucleotides long, are now termed short-interfering RNAs (siRNAs).

Subsequent genetic and biochemical analyses in several organisms took the shape of RNA silencing pathway. In vitro experiments in fly (*Drosophila melanogaster*) embryos demonstrated that long exogenous dsRNA is cleaved into siRNAs by Dicer, a dsRNA-specific RNaseIII-like endonuclease (Bernstein et al. 2001), which finally link dsRNA to the silencing. The characterization of ARGONAUTE1 (AGO1) protein solved the last piece of RNA silencing puzzle in plants (Baumberger and Baulcombe 2005; Eamens et al. 2008): that is, following the formation of dsRNA from single-stranded sense RNA by RNA-dependent RNA polymerase (RdRP), a Dicer-like (DCL) protein recognize and process that dsRNA into different classes of siRNAs from which one strand is then

incorporated into the RNA-induced silencing complex (RISC) containing the Argonaute (AGO) protein and this guides the cleavage of a target RNA with a complementary nucleotide sequence.

We now have a much greater understanding of the gene silencing pathways in plants and their crucial roles in a variety of biological regulation processes, such as development, plant defense against invading viral nucleic acids, and epigenetic modifications. siRNA directed target repression can occur in several levels including decreased RNA stability, lowered translational efficiency, and repression of chromatin modifications in the nucleus, which leads to transcriptional gene silencing (TGS) (Eamens et al. 2008). In addition, recent data demonstrate that plant siRNA plays a role in systemic silencing as a mobile signal, so that they can have effects over a long distance (Chitwood and Timmermans 2010; Dunoyer et al. 2010; Molnar et al. 2010). In addition to siRNA, another major class of small RNA molecules called microRNAs (miRNAs) have also been discovered in plants and proved to be an important negative regulator of gene expression (Voinnet 2009b). These gene silencing phenomena that are mediated by small non-coding RNAs are collectively called RNA silencing (Eamens et al. 2008).

RNA SILENCING PATHWAY IN PLANTS

Plants exhibit surprisingly diverse classes of small RNAs and the proteins that generate them. The currently known plant small RNAs can be categorized into two major groups based on differences in origin, biogenesis and mode of action: siRNA from long perfectly dsRNA precursors and miRNAs derived from single-stranded RNA transcripts (transcribed from MIR genes) with imperfectly fold-back stem-loop structures. Endogenously expressed siRNAs (endo-siRNAs) can be further divided into several secondary classifications, including: heterochromatic siRNAs (hc-siRNAs), secondary siRNAs and natural antisense transcript siRNAs (NAT-siRNAs) (Axtell 2013; Ghildiyal and Zamore 2009). Common features of all small RNAs are that members of the DCL family are recruited to process long dsRNAs into ~20-30 nucleotides, and one strand of the resulting small RNA duplexes are thereafter incorporated into AGO family proteins to hybridize with their complementary targets, functioning in a suppressive manner (Eamens et al. 2008). *Arabidopsis thaliana* encodes six RdRP enzymes (RDR), four DCL proteins and ten AGOs, with both unique and redundant functions. Each small RNA group recruits consistent and unique sets of RDR, DCL, and AGO family members for their distinct modes of biogenesis and function, which are also known to be conserved and to remain distinct from one another in multiple plant species.

The exo-siRNA pathway

Early examples of siRNAs were thought to be primarily exogenous in origin, derived directly from experimentally introduced dsRNAs or viral RNAs. Exogenous small

interfering RNAs (exo-siRNAs) generated from these invasive transcripts form a basis for antiviral defense.

In plants, exogenous sources of siRNAs are not limited to dsRNAs. Single-stranded RNAs (ssRNAs) from highly expressed transgenes or virus can be converted to dsRNA by RDR6/RDR1, members of the RdRP family that transcribe ssRNA from a RNA substrate (Figure 1.1) (Voinnet 2008; Qu et al. 2005). The dsRNA intermediary is then recognized and processed by a Dicer enzyme to produce two sizes of siRNAs: 21 and 24 nucleotides (Hamilton et al. 2002; Tang et al. 2003). The 21-nt siRNAs are processed by DCL4, but DCL2 is able to substitute and produce 22-nt siRNAs in the absence of DCL4 (Gascioli et al. 2005; Deleris et al. 2006; Dunoyer et al. 2007; Xie et al. 2004). The DCL4-produced 21-mers typically bind AGO1 and guide target transcript cleavage. The 24-nt siRNAs are generated by DCL3 and typically incorporated into AGO4 for the formation of repressive chromatin (Chan 2008).

The endo-siRNA pathway

hc-siRNA

Heterochromatic siRNAs are mostly ~24 nucleotides in size, generated from intergenic and/or repetitive genomic regions and promote the formation of repressive chromatin modifications (DNA methylation and histone modification) at homologous loci (Matzke et al. 2009; Law and Jacobsen 2010).

The biogenesis of heterochromatic siRNA requires specific members of the RDR, DCL, and AGO gene families. In the current model, plant-specific PolIV generates an aberrant RNA from methylated DNA templates (Figure 1.2) (Mosher et al. 2008). This aberrant RNA is then converted into dsRNA by RDR2 and subsequently processed into 24nt siRNA duplexes by DCL3 (Xie et al. 2004; Haag et al. 2012). The siRNA duplexes are methylated by HEN1 and preferentially bind AGO4-clade AGOs to guide RNA-directed DNA methylation (RdDM) that is catalyzed by the de novo cytosine methyltransferase DRM2 (Law and Jacobsen 2010). A second plant-specific RNA polymerase, Pol V is thought to be critical for the actual sequence-specific DNA methylation (Wierzbicki et al. 2009). Additional interacting partners are also recruited in steps downstream of siRNA biogenesis (Kanno et al. 2004; Kanno et al. 2008; Ausin et al. 2009; Bies-Etheve et al. 2009; He et al. 2009).

The ~24 nt heterochromatic siRNAs, which are easily distinguishable from the other classes of endogenous plant small RNAs, are clearly conserved in multiple species (Axtell 2013). In fact, small RNA sequencing experiments suggest that 24-nt sRNAs are the most abundant size in numerous flowering plants (Axtell 2013). However, it seems that only the heterochromatic siRNA pathway is evolutionarily ancient; in contrast, individual heterochromatic siRNA loci appear to have high rates of birth and death possibly due to the rapid changes of transposable elements in the plant genome during evolution (Axtell 2013; Ma et al. 2010).

The function of hc-siRNAs is largely to maintain genome integrity and stability, by assembly and maintenance of silent chromatin on transposable elements, which ultimately protect the genome from internal and external threats (Fei et al. 2013).

Secondary siRNA

Secondary siRNAs derive from a dsRNA precursor whose synthesis depends on the activity of one or more upstream small RNAs. In recent years, a novel class of phased secondary siRNA termed *trans*-acting small interfering RNAs (tasiRNAs), has been identified and intensively studied for their biogenesis and functions (Peragine et al. 2004; Hunter et al. 2006; Montgomery et al. 2008a; Williams et al. 2005; Xie et al. 2005; Yoshikawa et al. 2005; Garcia et al. 2006; Adenot et al. 2006; Allen et al. 2005; Vazquez et al. 2004). The generation of tasiRNAs involves components from both siRNA and miRNA pathways, including AGO1, DCL1, HEN1, HYPOPLASTIC LEAVES1 (HYL1), suppressor of gene silencing 3 (SGS3), and RDR6 (Peragine et al. 2004; Allen et al. 2005; Vazquez et al. 2004). Like miRNAs, but in contrast to other classes of endogenous siRNAs, tasiRNAs target genes that have little to no overall similarity with the genes from which the siRNAs are derived. However, unlike miRNAs, their production requires RDR6 and SGS3 activity, which are originally associated with viral defense (virus induced gene silencing) and transgene silencing (posttranscriptional gene silencing) (Mourrain et al. 2000). The identification of tasiRNAs establishes a link between the miRNA and siRNA pathways and also raises the possibility that numerous endogenous genes may be regulated by this special subgroup of endogenous siRNAs.

The generation of tasiRNAs is triggered by miRNA-guided cleavage of a tasiRNA-generating (TAS) gene derived transcript. One of the 3' or 5' cleavage products is converted to dsRNA by RDR6 and subsequently processed by DCL4 into 21-nt siRNAs that are phased with regard to the miRNA cleavage site (Figure 1.2) (Peragine et al. 2004; Allen et al. 2005; Xie et al. 2005; Yoshikawa et al. 2005). One strand of the tasiRNA duplex is selectively assembled into RNA-induced silencing complex (RISC) to catalyze the cleavage or repress the translation of complementary target mRNAs. *Arabidopsis thaliana* contains eight tasiRNA-generating (TAS) loci belonging to four families. *TAS1*, *TAS2*, *TAS4* only require one miRNA binding site upstream of the tasiRNA-generating region; while *TAS3* requires two, one upstream and one downstream of the tasiRNA-generating region (Fig. 1). These two different classes each requires unique components and unique RNA structural features for their biogenesis of tasiRNAs.

In the current model, both *TAS1* and *TAS2* require the miR173-AGO1 complex to guide the cleavage of transcripts for tasiRNA formation (Montgomery et al. 2008b). The synthesis of the complementary strand is directed by RDR6 from the 3' poly(A) tail towards the cleavage site, thus it is unknown how the miR173-AGO1 complex transmits a signal to recruit RDR6 distally. The 5' RNA fragments may subsequently undergo degradation by exoribonucleases, as no small RNAs have been identified from these sequences (Souret et al. 2004). *TAS4* is a target of miR828, yet it shares the similarity with *TAS1/2* in the way TAS transcripts are processed (Luo et al. 2012; Rajagopalan et al. 2006).

Unlike the *TAS1/2/4*, *TAS3* tasiRNAs originate from sequences between two miR390 binding sites. In these transcripts, miR390 guides cleavage on the 3' side of the

tasiRNA-generating region to set the processing register. The 5' miR390 complementary site is non-cleavable, which seems to be conserved among flowering plants except moss or pine (Axtell et al. 2006). In addition, binding of the AGO7-miR390 complex to the upstream, non-cleavable miR390 target site is essential for the production of tasiRNAs (Montgomery et al. 2008a; Axtell et al. 2006). In contrast, the 3' miR390 target site is cleaved across plant species and is not dependent on AGO7 or miR390 specifically, only that a functional cleavage site is present (Axtell et al. 2006). Notably, miR390 is unique from other miRNAs and is specifically associated with AGO7 (Montgomery et al. 2008a).

TAS2 has been predicted to produce at least one tasiRNA targeting two clusters of pentatricopeptide repeat gene transcripts (PPRs) in *Arabidopsis* (Allen et al. 2005). Interestingly, this *TAS2*-derived tasiRNA can itself initiate secondary siRNA production on these same PPR mRNAs. Although the *TAS1* loci do not have extended sequence similarity to *TAS2*, all three loci (*TAS1a*, *TAS1b*, and *TAS1c*) produce identical or very closely related tasiRNAs that also target a group of *PPR* genes as well as another group of genes of unknown function (Peragine et al. 2004; Vazquez et al. 2004; Yoshikawa et al. 2005). Currently, there is no other targets that has been identified for *TAS1* and *TAS2* tasiRNAs. In contrast to *TAS1/2*, one of the *TAS4*-derived tasiRNAs specifically targets a group of MYB transcription factors including *PAP1*, *PAP2*, and *MYB113* which are involved in the regulation of the anthocyanin biosynthesis pathway (Rajagopalan et al. 2006). All *TAS3* transcripts identified to date has been predicted to produce at least one tasiRNA (tasiARFs) that targets a family of Auxin Response Factor (*ARF1*, *ARF2*, *ARF3*, and *ARF4*) (Garcia et

al. 2006; Fahlgren et al. 2006; Adenot et al. 2006; Williams et al. 2005; Axtell et al. 2006; Allen et al. 2005; Hunter et al. 2006; Marin et al. 2010).

NAT-siRNA

Natural antisense transcript siRNAs (NAT-siRNAs) are a third subset of plant endo-siRNAs derived from two convergently transcribed mRNAs. NAT-siRNAs are further categorized into two subgroups: *cis*-NAT-siRNAs that are transcribed from opposite strands of the same locus, and *trans*-NAT-siRNAs from genes that harbor regions of complementarity (Lapidot and Pilpel 2006; Jin et al. 2008; Zhang et al. 2013). Typically, one transcript is produced constitutively, whereas the expression of complementary RNA only occurs in response to environmental stress, such as bacterial pathogen infection (Katiyar-Agarwal et al. 2006; Borsani et al. 2005).

Unlike other siRNAs, NAT-siRNAs have variable RDR and DCL requirements for biogenesis, variable sRNA size distributions and can regulate gene expression through distinct mechanisms (Zhang et al. 2013). The most straightforward pathway starts with the formation of dsRNAs by hybridization of two complementary RNAs (Figure 1.2). The overlap region between the two transcripts is then processed into NAT-siRNAs by the endonuclease activity of DCI2 and/or DCI1 (Katiyar-Agarwal et al. 2006; Borsani et al. 2005; Ron et al. 2010). NAT-siRNAs function mainly at the posttranscriptional level by cleavage of one transcript of the pair, and in some cases, trigger the DCI1-dependent production of 21-nt secondary siRNAs (Borsani et al. 2005). The accumulation of

secondary NAT-siRNAs is dependent on the function of RDR2/RDR6 and SGS3 (Zhang and Trudeau 2008; Zhang et al. 2013).

Genome-wide analysis indicates that up to 9% of *Arabidopsis* genes are overlapping and can potentially generate cis-NAT-siRNAs (Zhang et al. 2013). In *Arabidopsis*, three cis-NAT-siRNAs have been functionally analyzed: nat-siRNASRO5 induced by salt stress (Borsani et al. 2005), nat-siRNAATGB2 that accumulates in response to infection with avirulent bacteria (Katiyar-Agarwal et al. 2006), and nat-siRNA from the *KPL-AR114* locus (Ron et al. 2010). However, there are still important unanswered questions regarding to the biogenesis and functions of this group of siRNAs, which clearly need further exploration.

The miRNA pathway

miRNAs are the second most abundant class of plant small RNAs. Most plants encode more than 100 miRNA genes (MIR), mainly found in intergenic regions throughout the genome (Rogers and Chen 2013). Several miRNA families are conserved over long evolutionary distances, indicating their very ancient origin. These include miR156, miR160, miR319, miR390, all of which regulate ancestral transcription factors that impact the development, growth, and physiology of plants (Garcia 2008). However, the majority of miRNAs present in any given plant species are unique to that species or exist only between closely related species (Voinnet 2009a). miRNAs are key negative regulators of eukaryotic gene expression, and are widely believed to control a vast array of biological processes in

plants and animals, ranging from housekeeping functions to responses to biotic/abiotic stress.

In plants, RNA Pol II produces capped and polyadenylated primary miRNAs (pri-miRNAs) that contain an imperfect, self-complementary foldback region (Figure 1.2). In the nucleus, the pri-miRNA transcript is recognized and cleaved by DCL1, with the help of HYPONASTIC LEAVES1 (HYL1), to produce ~70-nucleotide precursor miRNA (pre-miRNA) molecule. Liberating mature miRNA from the pre-miRNA stem loop structure requires a second cleavage step which is again directed by DCL1 and its partner HYL1 (Vazquez et al. 2008). The two-nucleotide 3' overhangs of the released miRNA/miRNA* duplex are 2'-O-methylated by the methyltransferase HUA ENHANCER1 (HEN1) (Yang et al. 2006; Yu et al. 2005). The miRNA duplex is then transported to the cytoplasm, where the mature single-stranded miRNA is loaded onto AGO1, to direct the repressive regulation of complementary targets. Loss of Dicer or miRNA-associated Argonaute proteins almost always result in severe developmental abnormalities in both plants and animals. In *Arabidopsis* species, *dcl1* mutants exhibit abnormal embryogenesis, indicating that miRNA-mediated regulation is integral to pathways governing growth and development. (Nodine and Bartel 2010).

In flies and mammals, most miRNAs hybridize with their targets through a region of 6-8 nucleotides at the 5' end of the miRNA called the 'seed region', and chiefly direct translational repression of their targets (Lewis et al. 2003; Lewis et al. 2005). In contrast, the activity of plant miRNAs, which are highly complementary to targets throughout their length, was previously suggested to be chiefly accomplished via transcript cleavage

(Rhoades et al. 2002). Such extensive miRNA/target complementarity is considered the norm in plants, which occur most commonly in protein-coding regions of mRNAs (Llave et al. 2002; Rhoades et al. 2002; Tang et al. 2003). However, several examples have been known where a plant miRNA regulates target's steady-state protein level in the absence of noticeable changes in mRNA level, suggesting that plant miRNAs can also block translation (Aukerman and Sakai 2003; Chen 2004; Gandikota et al. 2007; Brodersen et al. 2008; Dugas and Bartel 2008; Beauclair et al. 2010). Although the mechanism of translational repression in plants still remains unknown, genetic studies have begun to uncover specific effectors of this pathway (Brodersen et al. 2008; Yang et al. 2012). Further analyses have clearly demonstrated that most plant miRNAs repress their targets via some combination of cleavage and translational repression, and that AGO catalyzed slicing often plays a key role (Carbonell et al. 2012; Yang et al. 2012). Nonetheless, translational inhibition is not a simple "back-up" system for slicing, and understanding the balance between these two modes of miRNA action (cleavage versus translational repression) awaits further investigation.

Plant miRNAs are predominately 21 nucleotides in size with a 5' U, indicating their DCL1-dependent processing and activities with AGO1 for posttranscriptional gene silencing (PTGS) of expressed target genes. (Cuperus et al. 2010; Chen et al. 2010). However, DCL1 generates alternative product sizes from foldback precursors containing asymmetric bulges (Kurihara and Watanabe 2004; Chen et al. 2010; Cuperus et al. 2010; Manavella et al. 2012). Differences in miRNA size may make them functionally distinct, for example, 22-nt miRNAs play a key role in triggering the production of secondary

siRNAs from target transcripts (for details, see secondary siRNA part) (Manavella et al. 2012; Cuperus et al. 2010; Chen et al. 2010). The remaining DCLs may produce a small subset of total miRNAs (Rajagopalan et al. 2006; Vazquez et al. 2008; Amor et al. 2009). In *Arabidopsis* and rice, DCL3 generates long miRNAs 23 to 25 nucleotides in length, which enter the heterochromatic siRNA effector pathway after their production (Vazquez et al. 2008). Similar to hc-siRNA, DCL3-dependent long miRNAs are sorted to AGO4 and direct cytosine DNA methylation at both MIR and target loci (Wu et al. 2010).

RNAi DELIVERY IN PLANTS

Transgene-induced RNA silencing

Early examples of RNAi were achieved by transforming plants with constructs to produce antisense and sense over-expression (cosuppression) RNA (Figure 1.3) (Jorgensen et al. 2006). However, RNAi can be more efficiently induced by expressing a transgene that is made from an inverted repeat (IR) sequence of the target gene separated by a spacer sequence (hairpin structure) (Figure 1.3). hpRNA induced RNAi has been proven to be remarkably efficient and could be used to silence a wide selection of target genes: almost 100% of transgenic plants display gene silencing and the phenotype obtained could be similar to those counterpart full loss-of-function mutants (Kusaba 2004; Chuang and

Meyerowitz 2000; Wesley et al. 2001; Watson et al. 2005). In fact, the majority transgene-induced RNA silencing in plants is achieved using a hairpin transgene (Frizzi and Huang 2010). Using an intron-derived spacer sequence is thought to be beneficial for the stability of hpRNA and thus enhance the silencing efficiency (Wesley et al. 2001). If a fragment of promoter sequence is used in the IR, TGS accompanied by de novo methylation can be efficiently triggered to silence the target gene (Figure 1.3) (Jones et al. 2001; Kanno et al. 2004; Mette et al. 2000; Sijen et al. 2001; Huettel et al. 2007).

In recent years, another RNAi delivery method utilizing artificial tasiRNAs has been developed and gained even more significant attention because of its success in gene silencing in plants (Peragine et al. 2004; Hunter et al. 2006; Montgomery et al. 2008a; Williams et al. 2005; Xie et al. 2005; Yoshikawa et al. 2005; Garcia et al. 2006; Adenot et al. 2006; Allen et al. 2005; Vazquez et al. 2004). The generation of tasiRNAs is triggered by miRNA-guided cleavage of a TAS gene derived transcript, resulting in the production of 21 nucleotide siRNAs that are phased with regard to the miRNA cleavage site (Peragine et al. 2004; Allen et al. 2005; Xie et al. 2005; Yoshikawa et al. 2005). In an artificial tasiRNA inducing vector, the transgene is simply made from a modified TAS gene sequence by substituting a single or several copies of native siRNA with atasiRNAs targeting particular genes (Felippes and Weigel 2009; Montgomery et al. 2008b; Montgomery et al. 2008a; de la Luz Gutiérrez-Nava et al. 2008). Compared with hpRNA, the atasiRNA approach is more practical for stacking multiple functional small RNA sequences into a single construct and eliminating the possibility of off-target effects.

Another promising new approach for RNAi delivery in plants is to engineer the endogenous miRNA precursors to produce an artificial miRNA (amiRNA) that selectively inhibit the expression of target genes (Figure 1.3). In contrast to hpRNA that are processed in to a population of siRNAs having varying sequences, small RNAs from a miRNA locus have a strong strand bias and usually only one predominant small RNA is generated. When both sequences of the miRNA duplex are altered without changing the structural features, amiRNA of desired sequence would be preferentially accumulated at a high level. This approach has proven to be effective with improved targeting specificity in various plant species (Kim and Somers 2010; Schwab et al. 2006; Khraiwesh et al. 2008; Parizotto et al. 2004; Alvarez et al. 2006). The higher precision and strand-specificity made this strategy a good alternative for the delivery of RNAi in plants, but one drawback is that one single amiRNA might not work well in some instances.

Virus-induced gene silencing (VIGS)

An alternative way to trigger RNA silencing in plants is virus-induced gene silencing (VIGS), which employs the plant's ability to cope with viral RNAs (Figure 1.3) (Lu et al. 2003). The replication intermediate dsRNA of the viral genome or the foldbacks in single-stranded viral RNA can be processed into siRNAs by DCL proteins. A modified viral genome with plant sequences can then generate siRNAs targeting the endogenous plant genes (Robertson 2004; Burch-Smith et al. 2004). VIGS has the advantage that it does not require development of stable transformants, and thus rapidly generates

phenotypes (Burch-Smith et al. 2004). However, the current limitation of this method is that most reliable and effective VIGS vectors have a limited host range upon which the VIGS vector is based (Robertson 2004).

APPLICATION OF RNAi IN SOYBEAN

RNAi is now widely used as a homology-based gene silencing tool to knock down specific target genes in various plant species including Rice, Banana, Tomato etc. (Angaji et al. 2010; Lopez-Gomollon and Dalmay 2010). It has also been employed for gene function analysis and genetic improvement of soybean (Kasai and Kanazawa 2012). Some of those applications are briefly discussed below.

Metabolic engineering

Because soybean seeds are the most important product of soybean plants, one of the main focus of RNA silencing in soybean has been modifications to seed components. Metabolic pathways in developing seeds have been targeted to accumulate nutritional valuable metabolites or divert unwanted ones. For instance, the immunodominant soybean allergen Gly m Bd 30 K protein has been greatly reduced by RNA silencing in transgenic soybean seeds (Herman et al. 2003). In other examples, phytic acids (Nunes et al. 2006;

Shi et al. 2007) and saponin (Takagi et al. 2011) were reduced to achieve the desirable phenotypes with nutritional or commercial interests. Seed storage protein composition was also modified by two independent research groups (Schmidt et al. 2011; Kinney et al. 2001).

The fatty acid composition of soybean seed has also been modified to lower the polyunsaturated fatty acid content for more stable soybean oils. The first manifestation of this approach was done by introducing a transgene that transcribes sense RNA homologous to the *FAD2-1* gene, and seed oleic acid content was increased as a result of the cosuppression of the target gene (KINNEY 1996). Metabolic engineering of the fatty acid biosynthetic pathway has also targeted soybean $\Delta 15$ desaturase to produce very long-chain polyunsaturated fatty acids (Chen et al. 2006). A few years later, hpRNA-mediated gene silencing was adopted to down-regulate two key fatty acid desaturase genes. Wang et al. developed a silencing construct containing an inverted-repeat fragment of the *GmFAD2* gene and obtained high oleic acid content that ranged from 71.5 to 81.9% (Wang and Xu 2008). At approximately the same time, an hpRNA-based RNAi vector was designed to effectively silence the three active members of the soybean *FAD3* gene family (Flores et al. 2008). Due to the 318-nt conserved *FAD3* sequence used to generate the inverted repeats, the resulting transgenic soybeans exhibited a *FAD3* null-like phenotype (Flores et al. 2008). In one special case, an intron sequence was used as inverted repeats to successfully down-regulate the soybean fatty acid desaturase *FAD2-1A* (Wagner et al. 2011). In contrast to the previous belief that RNA silencing is a cytoplasmic event, this result implies that PTGS can take place in the nucleus.

Virus resistance

The effectiveness of RNAi technology in soybean was also used to develop resistance to diseases, especially those caused by viruses. Resistance to viruses involves the use of a transgene that encodes a virus-derived transcript with the purpose to trigger RNA silencing against the viral RNA (Goldbach et al. 2003). Using this approach, soybean plants resistant against two viruses have been reported to date, that is *Soybean mosaic virus* (Furutani et al. 2006; Furutani et al. 2007; Wang et al. 2001) and *Soybean dwarf virus* (Tougou et al. 2006; Tougou et al. 2007).

In addition to viruses, RNAi constructs have also been used effectively to provide resistance to nematodes in transgenic soybean plants. Transgenic soybeans were generated using inverted repeats targeting the major sperm protein (MSP) gene from *H. glycines*, in which the reproductive potential of *H. glycines* were significantly reduced by the accumulation of MSP-specific siRNAs (Steeves et al. 2006). Similar strategies were also employed by another group for the control of *H. glycines* (Li et al. 2010) or rootknot nematode (*Meloidogyne incognita*) (Ibrahim et al. 2011). In contrast, use of amiRNA to down-regulate the LRR-kinase gene failed to alter the soybean cyst nematode resistance (Melito et al. 2010).

Gene function analysis via VIGS

VIGS is often used for the functional analysis of plant genes for its ease to make constructs and deliver desired nucleic acids into plant cells. This approach does not involve a transformation process, thus allows characterization of phenotypes of essential genes. (Burch-Smith et al. 2004; Lu et al. 2003). Up to date, there are at least 11 RNA virus vectors and five DNA virus vectors available for gene silencing in plants (Kanazawa 2008; Kasai and Kanazawa 2012). Three vectors have been applied to soybean: *Bean pod mottle virus* vector (Zhang and Ghabrial 2006; Zhang et al. 2010; Kachroo et al. 2008; Fu et al. 2009; Zhang et al. 2009; Meyer et al. 2009; Pandey et al. 2011; Singh et al. 2011), *Cucumber mosaic virus* vector (Nagamatsu et al. 2007; Nagamatsu et al. 2009; Liu et al. 2010) and *Apple latent spherical virus* vector (Yamagishi and Yoshikawa 2009).

VIGS has been used to characterize the role of candidate genes of a specific pathway in soybean. For example, the gene function of putative *flavonoid 3'-hydroxylase* (*F3'H*) gene was established using the *Cucumber mosaic virus* based VIGS vector (Nagamatsu et al. 2007). Also, the VIGS of *chalcone synthase* (*CHS*) gene was linked to the yellow seed coat phenotype in soybean seeds using the same strategy (Nagamatsu et al. 2007). Similarly, a candidate gene *GmTFL1b* was successfully associated with the soybean *Dt1* locus through a VIGS experiment (Liu et al. 2010). In addition to the candidate gene approach, VIGS has also been exploited to identify genes involved in disease resistance, such as *Soybean mosaic virus*, *Bean pod mottle virus*, *Pseudomonas syringae*, *Phakopsora pachyrhizi* (Pandey et al. 2011; Fu et al. 2009; Singh et al. 2011; Meyer et al. 2009; Kachroo et al. 2008). Especially, Singh et al. demonstrated that silencing of the three *GmFAD3*

homologous enhances jasmonic acid accumulation and, thereby, susceptibility to *Bean pod mottle virus* and *Pseudomonas syringae* in soybean (Singh et al. 2011).

IMPROVEMENT OF SYSTEMS FOR RNAi INDUCTION IN SOYBEAN

Soybean (*Glycine max*) is one of the most important crops in the world due to its high seed protein and oil content. Products made of soybean represent a great source for human foods and livestock feeds. Studying gene functions in soybean could not only help to improve its nutritional value but also provide a valuable source for studies on physiology and biochemistry. However, due to highly duplicated genome regions and to a large number of gene families, exploring gene functions in this crop is considered to be particularly difficult (Shoemaker et al. 1996). Hence, RNA silencing has now become the preferred methodology for the advantages it holds over conventional strategies, especially when it is carried out in a complex genome like soybean. In order to improve RNAi-mediated gene silencing as a tool to analyze gene function and manipulate commercial traits in soybean, an economically important gene, *GmFAD3*, was chosen as a test model in the present study.

GmFAD3

The oxidative stability and nutritional value of soybean oil is largely determined by the degree of saturation of relative triacylglycerol fatty acids (Kinney et al. 2002; Pattee et al. 2002). Desaturation of the fatty acids takes place in both the plastidial membrane and the endoplasmic reticulum (ER) membrane through the catalytic activity of fatty acid desaturases (Ohlrogge and Browse 1995; Singh et al. 2002). So far, at least 29 full-length soybean desaturase genes were identified, clustering into nine subfamilies (Chi et al. 2011). In the polyunsaturated fatty acid synthesis pathway, omega-3 fatty acid desaturase (*FAD3*) is responsible to add the third double bond to the linoleic acid precursor (18:2) for the production of Linolenic acid (18:3).

Multiple plastid-targeted and microsomal omega-3 fatty acid desaturases have been identified in plants, but the microsomal forms are primarily responsible for the seed linolenic acid level (Yadav et al. 1993; Bilyeu et al. 2003). Three independent microsomal omega-3 acid desaturases (*GmFAD3A*, *GmFAD3B*, *GmFAD3C*) have been characterized and linked to low seed linolenic acid phenotype in soybean (Bilyeu et al. 2003). *GmFAD3A* (Glyma14g37350) was significantly upregulated in developing seeds and played a predominant role in determining the linolenic acid content of seed storage oil (Bilyeu et al. 2003; Bilyeu et al. 2005; Bilyeu et al. 2006; Bilyeu et al. 2011). *GmFAD3B* (Glyma02g39230) shares 94% sequence similarity with *GmFAD3A* in the coding regions, while *GmFAD3C* (Glyma18g06950) only contain 79% identical DNA sequence. The expression levels of these two genes remained relatively low in developing soybean seeds compared to *GmFAD3A* and accordingly have a less impact on seed linolenic acid levels (Bilyeu et al. 2003; Bilyeu et al. 2005; Bilyeu et al. 2006; Bilyeu et al. 2011).

One of the most important goals of oil quality breeding in soybean has been to lower its α -linolenic acid (18:3) content for improved oxidative stability and flavor to eliminate the need for hydrogenation. Inhibition of *FAD3* in soybeans reduces the level of unstable linolenic acid (18:3) and the resultant soybean oil can be directly used without hydrogenation. We choose *FAD3* as a test model for our gene silencing assays because it is an economically valuable gene with an easily assayed, quantifiable phenotype (Miquel 1992). Meanwhile, the three gene family members make it an ideal model for the study of RNAi-mediated gene silencing in a complex genome like soybean.

FIGURES

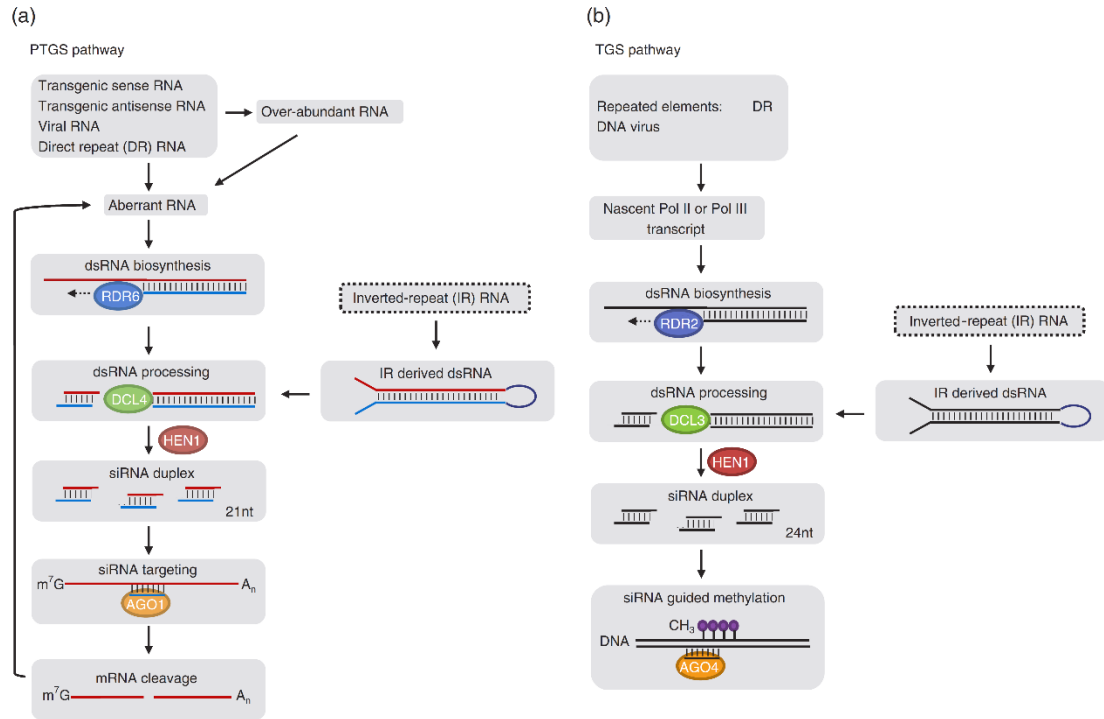


Figure 1.1 Exogenous siRNA pathways in plants. Courtesy from “**Tapping RNA silencing pathways for plant biotechnology**” (Frizzi and Huang 2010).

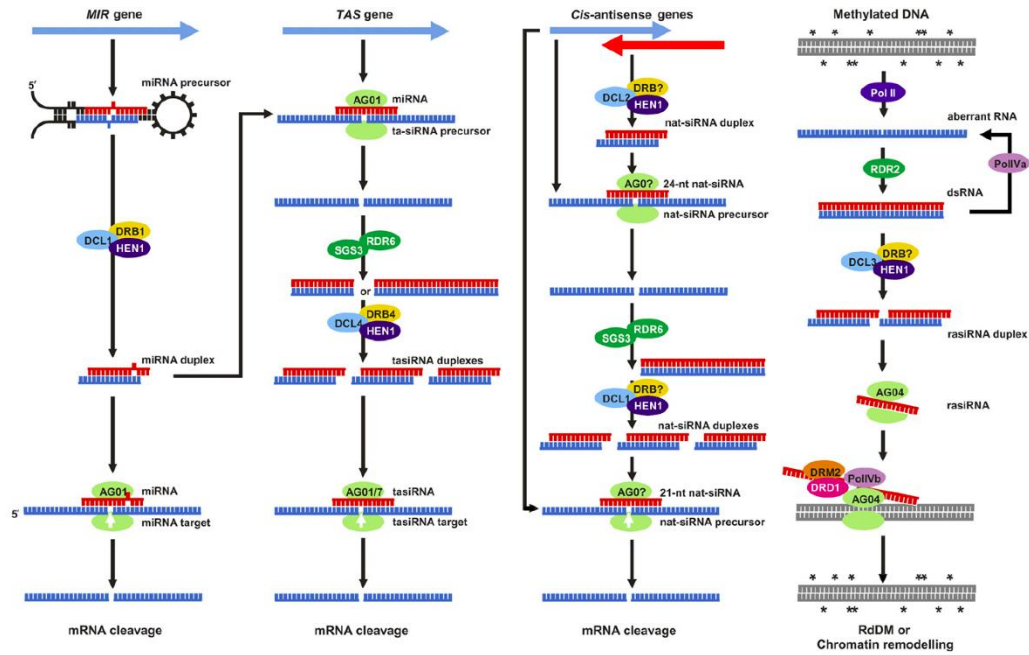


Figure 1.2 Endogenous small RNA pathways in plants. Courtesy from “RNA silencing in plants: yesterday, today, and tomorrow” (Eamens et al. 2008).

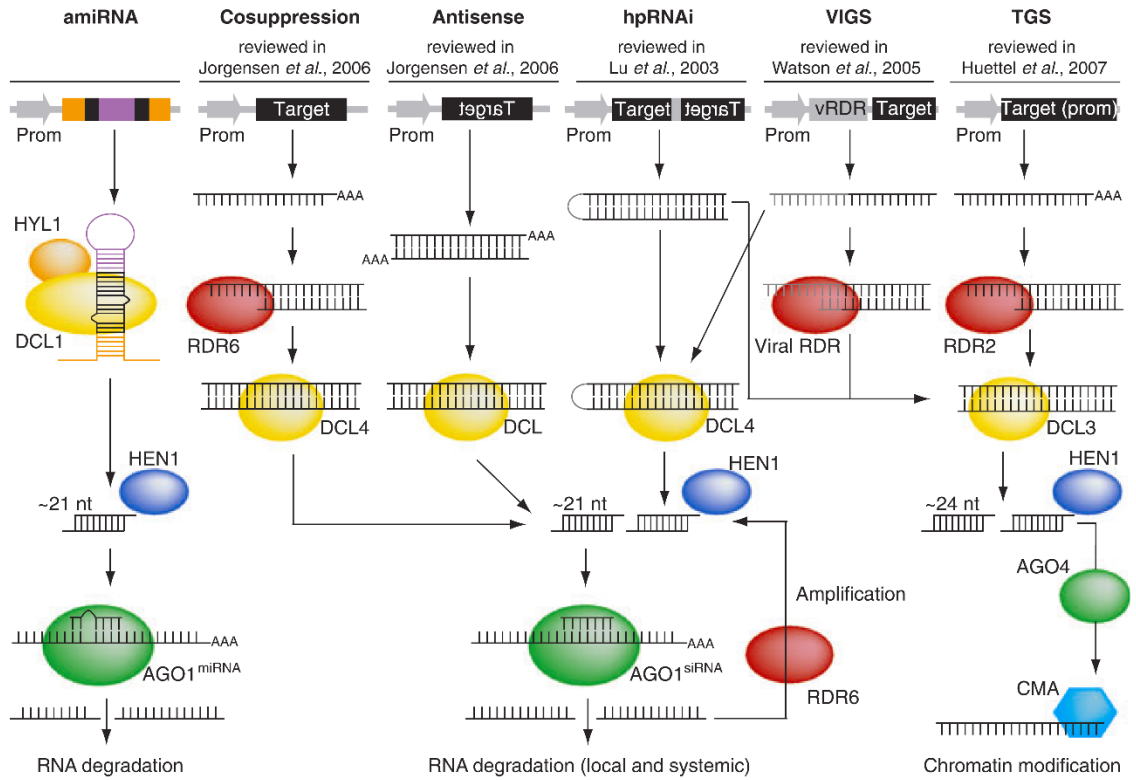


Figure 1.3 Comparison of transgene/virus-induced gene silencing pathways in plants. Courtesy from “Gene silencing in plants using artificial microRNAs and other small RNAs” (Ossowski et al. 2008).

REFERENCE

- Adenot X, Elmayan T, Laussergues D, Boutet S, Bouché N, Gasciolli V, Vaucheret H (2006) DRB4-Dependent *TAS3* *trans*-Acting siRNAs Control Leaf Morphology through AGO7. *Curr Biol* 16 (9):927-932
- Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-Directed Phasing during *Trans*-Acting siRNA Biogenesis in Plants. *Cell* 121 (2):207-221
- Alvarez JP, Pekker I, Goldshmidt A, Blum E, Amsellem Z, Eshed Y (2006) Endogenous and synthetic microRNAs stimulate simultaneous, efficient, and localized regulation of multiple targets in diverse species. *The Plant Cell Online* 18 (5):1134-1151
- Amor BB, Wirth S, Merchan F, Laporte P, d'Aubenton-Carafa Y, Hirsch J, Maizel A, Mallory A, Lucas A, Deragon JM (2009) Novel long non-protein coding RNAs involved in Arabidopsis differentiation and stress responses. *Genome Res* 19 (1):57-69
- Angaji SA, Hedayati SS, Hosein Poor R, Samad Poor S, Shiravi S, Madani S (2010) Application of RNA interference in plants. *Plant OMICS: Journal of Plant Molecular Biology & Omics* 3 (3)
- Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. *The Plant Cell Online* 15 (11):2730-2741
- Ausin I, Mockler TC, Chory J, Jacobsen SE (2009) IDN1 and IDN2 are required for de novo DNA methylation in Arabidopsis thaliana. *Nat Struct Mol Biol* 16 (12):1325-1327
- Axtell MJ (2013) Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol* 64:137-159. doi:10.1146/annurev-arplant-050312-120043
- Axtell MJ, Jan C, Rajagopalan R, Bartel DP (2006) A two-hit trigger for siRNA biogenesis in plants. *Cell* 127 (3):565-577

- Baulcombe DC (1996) Mechanisms of pathogen-derived resistance to viruses in transgenic plants. *The Plant cell* 8 (10):1833
- Baumberger N, Baulcombe D (2005) Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proceedings of the National Academy of Sciences of the United States of America* 102 (33):11928-11933
- Beauclair L, Yu A, Bouché N (2010) microRNA - directed cleavage and translational repression of the copper chaperone for superoxide dismutase mRNA in Arabidopsis. *The Plant Journal* 62 (3):454-462
- Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409 (6818):363-366
- Bies - Etheve N, Pontier D, Lahmy S, Picart C, Vega D, Cooke R, Lagrange T (2009) RNA - directed DNA methylation requires an AGO4 - interacting member of the SPT5 elongation factor family. *EMBO reports* 10 (6):649-654
- Bilyeu K, Gillman JD, LeRoy AR (2011) Novel Mutant Allele Combinations Produce Soybeans Containing 1% Linolenic Acid in the Seed Oil. *Crop Sci* 51 (1):259. doi:10.2135/cropsci2010.01.0044
- Bilyeu K, Palavalli L, Sleper D, Beuselinck P (2003) Three microsomal omega-3 fatty acid desaturase genes contribute to soybean linolenic acid levels. *Crop Sci* 43 (5):1833-1838
- Bilyeu K, Palavalli L, Sleper D, Beuselinck P (2005) Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. *Crop Sci* 45 (5):1830-1836
- Bilyeu K, Palavalli L, Sleper DA, Beuselinck P (2006) Molecular genetic resources for development of 1% linolenic acid soybeans. *Crop Sci* 46 (5):1913-1918
- Borsani O, Zhu J, Verslues PE, Sunkar R, Zhu J-K (2005) Endogenous siRNAs Derived from a Pair of Natural *cis*-Antisense Transcripts Regulate Salt Tolerance in *Arabidopsis*. *Cell* 123 (7):1279-1291

- Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. *Science* 320 (5880):1185-1190
- Burch - Smith TM, Anderson JC, Martin GB, Dinesh - Kumar SP (2004) Applications and advantages of virus - induced gene silencing for gene function studies in plants. *The Plant Journal* 39 (5):734-746
- Carbonell A, Fahlgren N, Garcia-Ruiz H, Gilbert KB, Montgomery TA, Nguyen T, Cuperus JT, Carrington JC (2012) Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. *The Plant Cell Online* 24 (9):3613-3629
- Chan SW (2008) Inputs and outputs for chromatin-targeted RNAi. *Trends Plant Sci* 13 (7):383-389
- Chen H-M, Chen L-T, Patel K, Li Y-H, Baulcombe DC, Wu S-H (2010) 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proceedings of the National Academy of Sciences* 107 (34):15269-15274
- Chen R, Matsui K, Ogawa M, Oe M, Ochiai M, Kawashima H, Sakuradani E, Shimizu S, Ishimoto M, Hayashi M (2006) Expression of $\Delta 6$, $\Delta 5$ desaturase and GLELO elongase genes from *Mortierella alpina* for production of arachidonic acid in soybean [*Glycine max*(L.) Merrill] seeds. *Plant Sci* 170 (2):399-406
- Chen X (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science* 303 (5666):2022-2025
- Chi X, Yang Q, Lu Y, Wang J, Zhang Q, Pan L, Chen M, He Y, Yu S (2011) Genome-wide analysis of fatty acid desaturases in soybean (*Glycine max*). *Plant Molecular Biology Reporter* 29 (4):769-783
- Chitwood DH, Timmermans MC (2010) Small RNAs are on the move. *Nature* 467 (7314):415-419
- Chuang C-F, Meyerowitz EM (2000) Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* 97 (9):4985-4990

- Cuperus JT, Carbonell A, Fahlgren N, Garcia-Ruiz H, Burke RT, Takeda A, Sullivan CM, Gilbert SD, Montgomery TA, Carrington JC (2010) Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. *Nat Struct Mol Biol* 17 (8):997-1003
- de la Luz Gutiérrez-Nava M, Aukerman MJ, Sakai H, Tingey SV, Williams RW (2008) Artificial trans-acting siRNAs confer consistent and effective gene silencing. *Plant Physiol* 147 (2):543-551
- Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* 313 (5783):68-71
- Dugas DV, Bartel B (2008) Sucrose induction of Arabidopsis miR398 represses two Cu/Zn superoxide dismutases. *Plant Mol Biol* 67 (4):403-417
- Dunoyer P, Himber C, Ruiz-Ferrer V, Alioua A, Voinnet O (2007) Intra- and intercellular RNA interference in Arabidopsis thaliana requires components of the microRNA and heterochromatic silencing pathways. *Nat Genet* 39 (7):848-856
- Dunoyer P, Schott G, Himber C, Meyer D, Takeda A, Carrington JC, Voinnet O (2010) Small RNA duplexes function as mobile silencing signals between plant cells. *Science* 328 (5980):912-916
- Eamens A, Wang MB, Smith NA, Waterhouse PM (2008) RNA silencing in plants: yesterday, today, and tomorrow. *Plant Physiol* 147 (2):456-468. doi:10.1104/pp.108.117275
- Fahlgren N, Montgomery TA, Howell MD, Allen E, Dvorak SK, Alexander AL, Carrington JC (2006) Regulation of *AUXIN RESPONSE FACTOR3* by *TAS3* ta-siRNA Affects Developmental Timing and Patterning in Arabidopsis. *Curr Biol* 16 (9):939-944
- Fei Q, Xia R, Meyers BC (2013) Phased, secondary, small interfering RNAs in posttranscriptional regulatory networks. *The Plant cell* 25 (7):2400-2415. doi:10.1105/tpc.113.114652

- Felippes FF, Weigel D (2009) Triggering the formation of tasiRNAs in *Arabidopsis thaliana*: the role of microRNA miR173. *EMBO reports* 10 (3):264-270
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391 (6669):806-811
- Flores T, Karpova O, Su X, Zeng P, Bilyeu K, Sleper DA, Nguyen HT, Zhang ZJ (2008) Silencing of GmFAD3 gene by siRNA leads to low α -linolenic acids (18: 3) of fad3-mutant phenotype in soybean [*Glycine max* (Merr.)]. *Transgenic Res* 17 (5):839-850
- Frizzi A, Huang S (2010) Tapping RNA silencing pathways for plant biotechnology. *Plant Biotechnol J* 8 (6):655-677
- Fu D-Q, Ghabrial S, Kachroo A (2009) GmRAR1 and GmSGT1 are required for basal, R gene-mediated and systemic acquired resistance in soybean. *Mol Plant-Microbe Interact* 22 (1):86-95
- Furutani N, Hidaka S, Kosaka Y, Shizukawa Y, Kanematsu S (2006) Coat protein gene-mediated resistance to soybean mosaic virus in transgenic soybean. *Breeding science* 56 (2):119-124
- Furutani N, Yamagishi N, Hidaka S, Shizukawa Y, Kanemastu S, Kosaka Y (2007) Soybean mosaic virus resistance in transgenic soybean caused by post-transcriptional gene silencing. *Breeding Science* 57 (2):123
- Gandikota M, Birkenbihl RP, Hohmann S, Cardon GH, Saedler H, Huijser P (2007) The miRNA156/157 recognition element in the 3' UTR of the *Arabidopsis* SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant J* 49 (4):683-693. doi:10.1111/j.1365-313X.2006.02983.x
- Garcia D A miRacle in plant development: role of microRNAs in cell differentiation and patterning. In: *Seminars in cell & developmental biology*, 2008. vol 6. Elsevier, pp 586-595
- Garcia D, Collier SA, Byrne ME, Martienssen RA (2006) Specification of Leaf Polarity in *Arabidopsis* via the *trans*-Acting siRNA Pathway. *Curr Biol* 16 (9):933-938

- Gascioli V, Mallory AC, Bartel DP, Vaucheret H (2005) Partially Redundant Functions of *Arabidopsis* DICER-like Enzymes and a Role for DCL4 in Producing *trans*-Acting siRNAs. *Curr Biol* 15 (16):1494-1500
- Ghildiyal M, Zamore PD (2009) Small silencing RNAs: an expanding universe. *Nat Rev Genet* 10 (2):94-108. doi:10.1038/nrg2504
- Goldbach R, Bucher E, Prins M (2003) Resistance mechanisms to plant viruses: an overview. *Virus Res* 92 (2):207-212
- Haag JR, Ream TS, Marasco M, Nicora CD, Norbeck AD, Pasa-Tolic L, Pikaard CS (2012) In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. *Mol Cell* 48 (5):811-818
- Hamilton A, Voinnet O, Chappell L, Baulcombe D (2002) Two classes of short interfering RNA in RNA silencing. *The EMBO Journal* 21 (17):4671-4679
- Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286 (5441):950-952
- He X-J, Hsu Y-F, Zhu S, Wierzbicki AT, Pontes O, Pikaard CS, Liu H-L, Wang C-S, Jin H, Zhu J-K (2009) An Effector of RNA-Directed DNA Methylation in *Arabidopsis* Is an ARGONAUTE 4-and RNA-Binding Protein. *Cell* 137 (3):498-508
- Herman EM, Helm RM, Jung R, Kinney AJ (2003) Genetic modification removes an immunodominant allergen from soybean. *Plant Physiol* 132 (1):36-43
- Huettel B, Kanno T, Daxinger L, Bucher E, van der Winden J, Matzke AJ, Matzke M (2007) RNA-directed DNA methylation mediated by DRD1 and Pol IVb: a versatile pathway for transcriptional gene silencing in plants. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1769 (5):358-374
- Hunter C, Willmann MR, Wu G, Yoshikawa M, de la Luz Gutiérrez-Nava M, Poethig SR (2006) Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in *Arabidopsis*. *Development* 133 (15):2973-2981

- Ibrahim HM, Alkharouf NW, Meyer SL, Aly MA, Gamal El-Din AEKY, Hussein EH, Matthews BF (2011) Post-transcriptional gene silencing of root-knot nematode in transformed soybean roots. *Exp Parasitol* 127 (1):90-99
- Jin H, Vacic V, Girke T, Lonardi S, Zhu J-K (2008) Small RNAs and the regulation of cis-natural antisense transcripts in *Arabidopsis*. *BMC Mol Biol* 9 (1):6
- Jones L, Ratcliff F, Baulcombe DC (2001) RNA-directed transcriptional gene silencing in plants can be inherited independently of the RNA trigger and requires Met1 for maintenance. *Curr Biol* 11 (10):747-757
- Jorgensen R, Doetsch N, Müller A, Que Q, Gendler K, Napoli C A paragenetic perspective on integration of RNA silencing into the epigenome and its role in the biology of higher plants. In: Cold Spring Harbor symposia on quantitative biology, 2006. Cold Spring Harbor Laboratory Press, pp 481-485
- Kachroo A, Fu D-Q, Havens W, Navarre D, Kachroo P, Ghabrial SA (2008) An oleic acid-mediated pathway induces constitutive defense signaling and enhanced resistance to multiple pathogens in soybean. *Mol Plant-Microbe Interact* 21 (5):564-575
- Kanazawa A (2008) RNA silencing manifested as visibly altered phenotypes in plants. *Plant Biotechnol* 25:423-435
- Kanno T, Bucher E, Daxinger L, Huettel B, Böhmendorfer G, Gregor W, Kreil DP, Matzke M, Matzke AJ (2008) A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. *Nat Genet* 40 (5):670-675
- Kanno T, Mette MF, Kreil DP, Aufsatz W, Matzke M, Matzke AJ (2004) Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. *Curr Biol* 14 (9):801-805
- Kasai M, Kanazawa A (2012) RNA silencing as a tool to uncover gene function and engineer novel traits in soybean. *Breeding science* 61 (5):468
- Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A, Zhu J-K, Staskawicz BJ, Jin H (2006) A pathogen-inducible endogenous siRNA in plant immunity. *Proceedings of the National Academy of Sciences* 103 (47):18002-18007

- Kennerdell JR, Carthew RW (1998) Use of dsRNA-Mediated Genetic Interference to Demonstrate that *frizzled* and *frizzled 2* Act in the Wingless Pathway. *Cell* 95 (7):1017-1026
- Khraiwesh B, Ossowski S, Weigel D, Reski R, Frank W (2008) Specific gene silencing by artificial MicroRNAs in *Physcomitrella patens*: an alternative to targeted gene knockouts. *Plant Physiol* 148 (2):684-693
- Kim J, Somers DE (2010) Rapid assessment of gene function in the circadian clock using artificial microRNA in *Arabidopsis mesophyll* protoplasts. *Plant Physiol* 154 (2):611-621
- Kinney A, Cahoon E, Hitz W (2002) Manipulating desaturase activities in transgenic crop plants. *Biochem Soc Trans* 30 (Pt 6):1099-1103
- KINNEY AJ (1996) Development of genetically engineered soybean oils for food applications. *Journal of Food Lipids* 3 (4):273-292
- Kinney AJ, Jung R, Herman EM (2001) Cosuppression of the α subunits of β -conglycinin in transgenic soybean seeds induces the formation of endoplasmic reticulum-derived protein bodies. *The Plant Cell Online* 13 (5):1165-1178
- Kurihara Y, Watanabe Y (2004) Arabidopsis micro-RNA biogenesis through Dicer-like 1 protein functions. *Proceedings of the National Academy of Sciences of the United States of America* 101 (34):12753-12758
- Kusaba M (2004) RNA interference in crop plants. *Curr Opin Biotechnol* 15 (2):139-143
- Lapidot M, Pilpel Y (2006) Genome - wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. *EMBO reports* 7 (12):1216-1222
- Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11 (3):204-220. doi:10.1038/nrg2719
- Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120 (1):15-20

- Lewis BP, Shih I-h, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. *Cell* 115 (7):787-798
- Li J, Todd TC, Oakley TR, Lee J, Trick HN (2010) Host-derived suppression of nematode reproductive and fitness genes decreases fecundity of *Heterodera glycines* Ichinohe. *Planta* 232 (3):775-785
- Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG (1993) Induction of a highly specific antiviral state in transgenic plants: implications for regulation of gene expression and virus resistance. *The Plant Cell Online* 5 (12):1749-1759
- Liu B, Watanabe S, Uchiyama T, Kong F, Kanazawa A, Xia Z, Nagamatsu A, Arai M, Yamada T, Kitamura K (2010) The soybean stem growth habit gene *Dt1* is an ortholog of *Arabidopsis* *TERMINAL FLOWER1*. *Plant Physiol* 153 (1):198-210
- Llave C, Xie Z, Kasschau KD, Carrington JC (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* 297 (5589):2053-2056
- Lopez-Gomollon S, Dalmay T (2010) Recent patents in RNA silencing in plants: constructs, methods and applications in plant biotechnology. *Recent patents on DNA & gene sequences* 4 (3):155-166
- Lu R, Martin-Hernandez AM, Peart JR, Malcuit I, Baulcombe DC (2003) Virus-induced gene silencing in plants. *Methods* 30 (4):296-303
- Luo Q-J, Mittal A, Jia F, Rock CD (2012) An autoregulatory feedback loop involving PAPI and TAS4 in response to sugars in *Arabidopsis*. *Plant Mol Biol* 80 (1):117-129
- Ma Z, Coruh C, Axtell MJ (2010) *Arabidopsis lyrata* small RNAs: transient MIRNA and small interfering RNA loci within the *Arabidopsis* genus. *The Plant Cell Online* 22 (4):1090-1103
- Manavella PA, Koenig D, Weigel D (2012) Plant secondary siRNA production determined by microRNA-duplex structure. *Proceedings of the National Academy of Sciences* 109 (7):2461-2466

- Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheret H, Nussaume L, Crespi MD, Maizel A (2010) miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *The Plant Cell Online* 22 (4):1104-1117
- Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ (2009) RNA-mediated chromatin-based silencing in plants. *Curr Opin Cell Biol* 21 (3):367-376
- Melito S, Heuberger AL, Cook D, Diers BW, MacGuidwin AE, Bent AF (2010) A nematode demographics assay in transgenic roots reveals no significant impacts of the *Rhg1* locus LRR-Kinase on soybean cyst nematode resistance. *BMC Plant Biol* 10 (1):104
- Mette M, Aufsatz W, Van der Winden J, Matzke M, Matzke A (2000) Transcriptional silencing and promoter methylation triggered by double - stranded RNA. *The EMBO Journal* 19 (19):5194-5201
- Meyer JD, Silva DC, Yang C, Pedley KF, Zhang C, van de Mortel M, Hill JH, Shoemaker RC, Abdelnoor RV, Whitham SA (2009) Identification and analyses of candidate genes for Rpp4-mediated resistance to Asian soybean rust in soybean. *Plant Physiology* 150 (1):295-307
- Miquel M (1992) Arabidopsis mutants deficient in polyunsaturated fatty acid synthesis. Biochemical and genetic characterization of a plant oleoyl-phosphatidylcholine desaturase. *J Biol Chem* 267 (3):1502-1509
- Molnar A, Melnyk CW, Bassett A, Hardcastle TJ, Dunn R, Baulcombe DC (2010) Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. *Science* 328 (5980):872-875
- Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008a) Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in *TAS3* Trans-Acting siRNA Formation. *Cell* 133 (1):128-141
- Montgomery TA, Yoo SJ, Fahlgren N, Gilbert SD, Howell MD, Sullivan CM, Alexander A, Nguyen G, Allen E, Ahn JH (2008b) AGO1-miR173 complex initiates phased

siRNA formation in plants. *Proceedings of the National Academy of Sciences* 105 (51):20055-20062

Mosher RA, Schwach F, Studholme D, Baulcombe DC (2008) PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. *Proceedings of the National Academy of Sciences* 105 (8):3145-3150

Mourrain P, Béclin C, Elmayan T, Feuerbach F, Godon C, Morel J-B, Jouette D, Lacombe A-M, Nikic S, Picault N (2000) *Arabidopsis* *SGS2* and *SGS3* Genes Are Required for Posttranscriptional Gene Silencing and Natural Virus Resistance. *Cell* 101 (5):533-542

Nagamatsu A, Masuta C, Matsuura H, Kitamura K, Abe J, Kanazawa A (2009) Down-regulation of flavonoid 3'-hydroxylase gene expression by virus-induced gene silencing in soybean reveals the presence of a threshold mRNA level associated with pigmentation in pubescence. *J Plant Physiol* 166 (1):32-39

Nagamatsu A, Masuta C, Senda M, Matsuura H, Kasai A, Hong JS, Kitamura K, Abe J, Kanazawa A (2007) Functional analysis of soybean genes involved in flavonoid biosynthesis by virus - induced gene silencing. *Plant Biotechnol J* 5 (6):778-790

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *The Plant Cell Online* 2 (4):279-289

Ngo H, Tschudi C, Gull K, Ullu E (1998) Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences* 95 (25):14687-14692

Nodine MD, Bartel DP (2010) MicroRNAs prevent precocious gene expression and enable pattern formation during plant embryogenesis. *Genes Dev* 24 (23):2678-2692

Nunes AC, Vianna GR, Cuneo F, Amaya-Farfán J, de Capdeville G, Rech EL, Aragão FJ (2006) RNAi-mediated silencing of the myo-inositol-1-phosphate synthase gene (*GmMIPS1*) in transgenic soybean inhibited seed development and reduced phytate content. *Planta* 224 (1):125-132

Ohlrogge J, Browse J (1995) Lipid biosynthesis. *The Plant cell* 7 (7):957

- Ossowski S, Schwab R, Weigel D (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant Journal* 53 (4):674-690
- Pandey AK, Yang C, Zhang C, Graham MA, Horstman HD, Lee Y, Zabolina OA, Hill JH, Pedley KF, Whitham SA (2011) Functional analysis of the Asian soybean rust resistance pathway mediated by Rpp2. *Mol Plant-Microbe Interact* 24 (2):194-206
- Parizotto EA, Dunoyer P, Rahm N, Himber C, Voinnet O (2004) In vivo investigation of the transcription, processing, endonucleolytic activity, and functional relevance of the spatial distribution of a plant miRNA. *Genes Dev* 18 (18):2237-2242
- Pattee HE, Isleib TG, Moore KM, Gorbet DW, Giesbrecht FG (2002) Effect of high-oleic trait and paste storage variables on sensory attribute stability of roasted peanuts. *Journal of agricultural and food chemistry* 50 (25):7366-7370
- Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS (2004) SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. *Genes Dev* 18 (19):2368-2379
- Qu F, Ye X, Hou G, Sato S, Clemente TE, Morris TJ (2005) RDR6 has a broad-spectrum but temperature-dependent antiviral defense role in *Nicotiana benthamiana*. *J Virol* 79 (24):15209-15217
- Rajagopalan R, Vaucheret H, Trejo J, Bartel DP (2006) A diverse and evolutionarily fluid set of microRNAs in *Arabidopsis thaliana*. *Genes Dev* 20 (24):3407-3425
- Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. *Cell* 110 (4):513-520
- Robertson D (2004) VIGS vectors for gene silencing: many targets, many tools. *Annu Rev Plant Biol* 55:495-519
- Rogers K, Chen X (2013) Biogenesis, turnover, and mode of action of plant microRNAs. *The Plant cell* 25 (7):2383-2399. doi:10.1105/tpc.113.113159

- Ron M, Saez MA, Williams LE, Fletcher JC, McCormick S (2010) Proper regulation of a sperm-specific cis-nat-siRNA is essential for double fertilization in *Arabidopsis*. *Genes Dev* 24 (10):1010-1021
- Schmidt MA, Barbazuk WB, Sandford M, May G, Song Z, Zhou W, Nikolau BJ, Herman EM (2011) Silencing of soybean seed storage proteins results in a rebalanced protein composition preserving seed protein content without major collateral changes in the metabolome and transcriptome. *Plant Physiol* 156 (1):330-345
- Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D (2006) Highly specific gene silencing by artificial microRNAs in *Arabidopsis*. *The Plant Cell Online* 18 (5):1121-1133
- Shi J, Wang H, Schellin K, Li B, Faller M, Stoop JM, Meeley RB, Ertl DS, Ranch JP, Glassman K (2007) Embryo-specific silencing of a transporter reduces phytic acid content of maize and soybean seeds. *Nat Biotechnol* 25 (8):930-937
- Shoemaker R, Polzin K, Labate J, Specht J, Brummer E, Olson T, Young N, Concibido V, Wilcox J, Tamulonis J (1996) Genome duplication in soybean (*Glycine subgenus soja*). *Genetics* 144 (1):329-338
- Sijen T, Vijn I, Rebocho A, van Blokland R, Roelofs D, Mol JN, Kooter JM (2001) Transcriptional and posttranscriptional gene silencing are mechanistically related. *Curr Biol* 11 (6):436-440
- Singh AK, Fu D-Q, El-Habbak M, Navarre D, Ghabrial S, Kachroo A (2011) Silencing genes encoding omega-3 fatty acid desaturase alters seed size and accumulation of bean pod mottle virus in soybean. *Mol Plant-Microbe Interact* 24 (4):506-515
- Singh SC, Sinha RP, Hader DP (2002) Role of lipids and fatty acids in stress tolerance in cyanobacteria. *Acta Protozool* 41 (4):297-308
- Souret FF, Kastenmayer JP, Green PJ (2004) AtXRN4 Degrades mRNA in *Arabidopsis* and Its Substrates Include Selected miRNA Targets. *Mol Cell* 15 (2):173-183
- Steeves RM, Todd TC, Essig JS, Trick HN (2006) Transgenic soybeans expressing siRNAs specific to a major sperm protein gene suppress *Heterodera glycines* reproduction. *Funct Plant Biol* 33 (11):991-999

- Tang G, Reinhart BJ, Bartel DP, Zamore PD (2003) A biochemical framework for RNA silencing in plants. *Genes Dev* 17 (1):49-63
- Tougou M, Furutani N, Yamagishi N, Shizukawa Y, Takahata Y, Hidaka S (2006) Development of resistant transgenic soybeans with inverted repeat-coat protein genes of soybean dwarf virus. *Plant Cell Rep* 25 (11):1213-1218
- Tougou M, Yamagishi N, Furutani N, Shizukawa Y, Takahata Y, Hidaka S (2007) Soybean dwarf virus-resistant transgenic soybeans with the sense coat protein gene. *Plant Cell Rep* 26 (11):1967-1975
- Vazquez F, Blevins T, Ailhas J, Boller T, Meins F (2008) Evolution of Arabidopsis MIR genes generates novel microRNA classes. *Nucleic Acids Res* 36 (20):6429-6438
- Vazquez F, Vaucheret H, Rajagopalan R, Lepers C, Gascioli V, Mallory AC, Hilbert J-L, Bartel DP, Cr  t   P (2004) Endogenous *trans*-Acting siRNAs Regulate the Accumulation of *Arabidopsis* mRNAs. *Mol Cell* 16 (1):69-79
- Voinnet O (2008) Use, tolerance and avoidance of amplified RNA silencing by plants. *Trends Plant Sci* 13 (7):317-328
- Voinnet O (2009a) Origin, biogenesis, and activity of plant microRNAs. *Cell* 136 (4):669-687. doi:10.1016/j.cell.2009.01.046
- Voinnet O (2009b) Origin, biogenesis, and activity of plant microRNAs. *Cell* 136 (4):669-687
- Wagner N, Mroczka A, Roberts PD, Schreckengost W, Voelker T (2011) RNAi trigger fragment truncation attenuates soybean FAD2 - 1 transcript suppression and yields intermediate oil phenotypes. *Plant Biotechnol J* 9 (7):723-728
- Wang G, Xu Y (2008) Hypocotyl-based Agrobacterium-mediated transformation of soybean (*Glycine max*) and application for RNA interference. *Plant Cell Rep* 27 (7):1177-1184
- Wang X, Eggenberger AL, Nutter Jr FW, Hill JH (2001) Pathogen-derived transgenic resistance to soybean mosaic virus in soybean. *Mol Breed* 8 (2):119-127

- Waterhouse PM, Graham MW, Wang M-B (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proceedings of the National Academy of Sciences* 95 (23):13959-13964
- Watson JM, Fusaro AF, Wang M, Waterhouse PM (2005) RNA silencing platforms in plants. *FEBS Lett* 579 (26):5982-5987
- Wesley SV, Helliwell CA, Smith NA, Wang M, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA (2001) Construct design for efficient, effective and high - throughput gene silencing in plants. *The Plant Journal* 27 (6):581-590
- Wierzbicki AT, Ream TS, Haag JR, Pikaard CS (2009) RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet* 41 (5):630-634
- Williams L, Carles CC, Osmont KS, Fletcher JC (2005) A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the Arabidopsis ARF2, ARF3, and ARF4 genes. *Proceedings of the National Academy of Sciences of the United States of America* 102 (27):9703-9708
- Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, Qi Y (2010) DNA methylation mediated by a microRNA pathway. *Mol Cell* 38 (3):465-475
- Xie Z, Allen E, Wilken A, Carrington JC (2005) DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* 102 (36):12984-12989
- Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* 2 (5):e104
- Yadav NS, Wierzbicki A, Aegerter M, Caster CS, Perez-Grau L, Kinney AJ, Hitz WD, Booth Jr JR, Schweiger B, Stecca KL (1993) Cloning of Higher Plant [ω]-3 Fatty Acid Desaturases. *Plant Physiol* 103 (2):467-476
- Yamagishi N, Yoshikawa N (2009) Virus-induced gene silencing in soybean seeds and the emergence stage of soybean plants with Apple latent spherical virus vectors. *Plant Mol Biol* 71 (1-2):15-24

- Yang L, Wu G, Poethig RS (2012) Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. *Proceedings of the National Academy of Sciences* 109 (1):315-320
- Yang Z, Ebright YW, Yu B, Chen X (2006) HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res* 34 (2):667-675
- Yoshikawa M, Peragine A, Park MY, Poethig RS (2005) A pathway for the biogenesis of trans-acting siRNAs in Arabidopsis. *Genes Dev* 19 (18):2164-2175
- Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, Steward R, Chen X (2005) Methylation as a crucial step in plant microRNA biogenesis. *Science* 307 (5711):932-935
- Zhang C, Bradshaw JD, Whitham SA, Hill JH (2010) The development of an efficient multipurpose bean pod mottle virus viral vector set for foreign gene expression and RNA silencing. *Plant Physiol* 153 (1):52-65
- Zhang C, Ghabrial SA (2006) Development of **Bean pod mottle** virus-based vectors for stable protein expression and sequence-specific virus-induced gene silencing in soybean. *Virology* 344 (2):401-411
- Zhang C, Yang C, Whitham SA, Hill JH (2009) Development and use of an efficient DNA-based viral gene silencing vector for soybean. *Mol Plant-Microbe Interact* 22 (2):123-131
- Zhang D, Trudeau VL (2008) The XS domain of a plant specific SGS3 protein adopts a unique RNA recognition motif (RRM) fold. *Cell Cycle* 7 (14):2268-2270
- Zhang X, Lii Y, Wu Z, Polishko A, Zhang H, Chinnusamy V, Lonardi S, Zhu JK, Liu R, Jin H (2013) Mechanisms of small RNA generation from cis-NATs in response to environmental and developmental cues. *Molecular plant* 6 (3):704-715. doi:10.1093/mp/sst051

CHAPTER 2

**Knocking-down *Glycine max* omega 3 fatty acid desaturase (*FAD3*)
gene: A case study revealing siRNA-mediated gene silencing process in
crop plants**

ABSTRACT

Since the discovery of RNA silencing in the nineties, the implication and potential application of this new technology have been recognized. In the past few years, RNA silencing has gained significant attention because its success in genomic scale research and also in the genetic improvement of crop plants. In order to improve hpRNA-mediated gene silencing in soybean, the *GmFAD3* gene family was chosen as a test model. In this study, all three family members of *GmFAD3* were successfully silenced and the silencing phenotype was stably inherited. Silencing levels of *FAD3A*, *FAD3B* and *FAD3C* correlate to degrees of sequence homology between the inverted repeats (IR) of hpRNA and *GmFAD3* transcripts in the RNAi lines. siRNAs generated from the 318-bp IR were characterized and associated with the inferred cleavage sites on target transcripts. Small RNAs corresponding to the loop portion of the hairpin transcript were detected, implicating possible transitive self-silencing of the hairpin transgene. In contrast, much less RNAs were found outside of the target region, suggesting that transitivity along endogenous transcripts is prohibited by some inherent protective feature. Strikingly, transgenes in two of the three RNAi lines were heavily methylated, leading to a dramatic reduction of hpRNA-derived siRNAs. Small RNAs encoding part of the transgene promoter as well as the *bar* gene coding sequences were also detected by deep sequencing, but whether they induced the methylation of transgenes still need further exploration.

INTRODUCTION

Since the discovery of RNA silencing in the nineties, the implication and potential application of this new technology have been recognized. RNAi has advantages over classic mutagenesis not only because it has the ability of silencing multiple gene family members with one single RNAi-inducing construct but also because it has the potential capability to control the suppression in a regulated manner. In the past few years, RNA silencing has gained significant attention because its success in genomic scale research and also in the genetic improvement of crop plants (Frizzi and Huang 2010; Watson et al. 2005). For example, RNAi technology was used to suppress the caffeine synthase gene to create varieties of *Coffea* that produces natural coffee with reduced caffeine content (Ogita et al. 2003).

In plants, the RNAi pathway primarily deploys siRNAs for sequence-specific target mRNA degradation (Frizzi and Huang 2010). The delivery of siRNAs can be achieved by expressing a transgene that is made from an inverted repeat sequence of a target gene separated by an intron as a spacer (hairpin structure). The resulting 21 nucleotide-long small RNA molecules with sequence complementarity to the target mRNA then direct either degradation or translational repression of those designated transcripts. hpRNA induced RNAi has been proven to be remarkably efficient and could be used to silence a wide selection of target genes. : almost 100% of transgenic plants display gene silencing and the phenotype obtained could be similar to those counterpart full loss-of-function

mutants (Kusaba 2004). However, most of the studies until now have been done in model plants with relatively simple genomes and the silencing efficiency and specificity of siRNA-mediated gene silencing have not been well characterized in crop plants.

Soybean is one of the most important crops in the world due to its high seed protein and oil content. Soybean oil is used extensively in the food industry and represented 57% of the world's oilseed production in 2012 (<http://soystats.com/2012>). Commodity soybean oil typically contains about 7-10% of linolenic acid (18:3), which is undesirable for many food applications for its oxidative instability (Liu and White 1992). While chemical hydrogenation has been employed to reduce the amount of linolenic acid to improve the quality of soybean oil, the process also created undesirable *trans*-fats that have been linked to many health problems in humans, particularly coronary heart disease (Hu et al. 1997; Ascherio and Willett 1997). Therefore, one of the most important goals of oil quality breeding in soybean has been to lower its linolenic acid content for improved oxidative stability and flavor to eliminate the need for hydrogenation. However, due to highly duplicated genome regions and to a large number of gene families, exploring gene functions and improve commercial traits in soybean is considered to be particularly difficult (Shoemaker et al. 1996). Thus, RNAi-mediated gene silencing has become the technology of choice for the advantages it holds over conventional strategies, especially when it is carried out in a complex genome like soybean.

Linolenic acid is produced from linoleic acid precursors (18:2) under the catalytic activity of omega-3 fatty acid desaturase (*FAD3*) in the polyunsaturated fatty acid synthesis pathway. Thus, inhibition of *FAD3* in soybeans reduces the level of unstable linolenic acid

and the resultant soybean oil can be directly used without hydrogenation. Multiple plastid-targeted and microsomal omega-3 fatty acid desaturases have been identified in plants, but the microsomal forms are primarily responsible for the seed linolenic acid level (Yadav et al. 1993; Bilyeu et al. 2003). Three distinct microsomal omega-3 acid desaturases (*GmFAD3A*, *GmFAD3B*, *GmFAD3C*) have been characterized and linked to low seed linolenic acid phenotype in soybean using candidate gene based approaches (Bilyeu et al. 2003). *GmFAD3A* (Glyma14g37350) has the highest expression level of the three homologs in developing seeds and has been shown to be the major contributor to seed linolenic acid levels (Bilyeu et al. 2003; Bilyeu et al. 2005; Bilyeu et al. 2006; Bilyeu et al. 2011). *GmFAD3B* (Glyma02g39230) shares 94% sequence similarity with *GmFAD3A* in the coding regions, while *GmFAD3C* (Glyma18g06950) only contain 79% identical DNA sequence. These two genes are much less expressed in developing soybean seeds compared to *GmFAD3A* and accordingly have a less impact on seed linolenic acid levels (Bilyeu et al. 2003; Bilyeu et al. 2005; Bilyeu et al. 2006; Bilyeu et al. 2011). We choose *FAD3* as a test model for our gene silencing assays because it is an economically valuable gene with an easily assayed, quantifiable phenotype. Meanwhile, the three gene family members make it an ideal model to test the efficacy and specificity of RNAi mediated silencing of gene families in soybean.

In our lab's previous research, a hpRNA-based RNAi vector pMUFAD was designed to effectively silence the three active members of soybean *FAD3* gene family (Flores et al. 2008). A 318-bp highly conserved nucleotide sequence representing a domain common among family members was used for the development of inverted repeats (IR),

separated by a spacer fragment derived from the intron of the rice waxy-a gene to form the hair-pin structure. A high level of silencing was achieved by transgene produced siRNAs, which led to a significant reduction of linolenic acid content in the seed oil. However, variations were detected in the down-regulated linolenic acid level between different RNAi lines, ranging from 1.2% to 3.6% in the T₃ homozygous seeds (Flores et al. 2008). Further investigation will then be needed to find out the possible molecular basis responsible for this phenomenon. Moreover, the relatively long inverted repeats used to generate RNAi by hpRNA may suffer from the off-target effect (Jackson et al. 2003). Furthermore, details about the complexity of RNA silencing in stably transformed soybean plants derived from *Agrobacterium*-mediated T-DNA transfer are still elusive. Thus, more effort could be made for the optimization of parameters in RNAi approach to generate strong, consistent and heritable soybean events with desired trait modifications.

This research is directed at improving RNAi technology as a tool to analyze gene function and manipulate commercial traits in soybean. To fully capitalize on the potential of RNAi, the endogenous soybean gene family *GmFAD3* was chosen as test model gene. The objectives of this work were: 1) to determine the heritability of the RNAi phenotype in stably transformed soybean; 2) to characterize the specificity of hpRNA mediated RNAi; and 3) to investigate the potential molecular basis for variations in RNAi silencing efficiency in different soybean events.

MATERIALS AND METHODS

Plant material and growth conditions

T₃ transgenic soybean plants were obtained from our lab's previous work (Flores et al., 2008) and grown until the T₅ generation. All soybean were grown on Pro-mix soil (SunGro, Agawam, MA) in 13-liter pots in a greenhouse under controlled-environmental conditions at 23-26°C with supplemental 50-90 Klux day light intensity and 12/12 h photoperiod from late May to early November or a 16/8 h photoperiod during the rest seasons. Plants were fertilized once with Osmocote 14-14-14 (Hummert International, Earth City, MO) at the time of planting and watered as needed.

Fatty acid analysis

The fatty acid profiles of dry mature soybean seeds from transgenic and wild type control samples were examined by a gas chromatography (GC) method as previously described (Beuselinck et al. 2006). A bulk sample of 5 seeds from each plant was crushed in an envelope and used as samples for fatty acid determination. For each transgenic soybean line, seeds from three plants were individually analyzed. The individual fatty acid

contents of palmitic, stearic, oleic, linoleic, and linolenic acids are presented as a proportion of total fatty acids in the extracted oil.

qRT-PCR

Mid-mature soybean seeds of a transgenic plant were collected and immediately frozen in liquid nitrogen, then stored at -80 for later use. Total RNA from each seed was extracted with TRIzol reagent (Invitrogen) and purified with a DNA-Free RNA kit (Zymo Research, Irvine, CA) to remove genomic DNA contamination. First-strand cDNA was synthesized from 500ng of the DNase-treated RNA using iScriptTM Reverse Transcription Supermix (Bio-Rad, Hercules, CA). The resulting cDNA was diluted to a final concentration of 10ng/ul for qRT-PCR analyses.

Real-time quantitative PCR was performed in triplicate biological and technical replications on an CFX-96TM Real-Time system (Bio-Rad, Hercules, CA) with the recommended settings for SYBR Green. Each reaction contained 2ul diluted cDNA, 10 μ M of each specific primer, and 10 μ l of 2x SsoAdvancedTM Universal SYBR[®] Green Supermix (Bio-Rad, Hercules, CA) in a final volume of 20 μ l. Genomic DNA and other contamination were monitored by no-template and no-RT controls. A standard curve was generated from pooled cDNAs to determine the PCR efficiency of each primer pair. The following PCR program was used for all PCR reactions: 95°C for 30 s, followed by 35

cycles of 10s denaturation at 95°C, 30s annealing and extension at 60°C. Amplification specificity was verified by melting curve analysis at the end of PCR.

Templates were normalized for differences in cDNA amount using *CONS7* amplification levels. Data were analyzed with BioRad CFX Manager™ 2.0 Software (BioRad, Hercules, CA). The comparative threshold cycle method ($\Delta\Delta C_t$) was used to determine relative transcript abundance levels. Sequences of applied primers are listed in Table 2.1.

Bisulfite Sequencing

Genomic DNA was isolated from mid-mature seeds of *FAD3* homozygous lines using the CTAB method (lab protocol) and further purified by a Genomic DNA Clean-up Kit (Zymo Research, Irvine, CA). 700ng DNA was bisulfite modified in duplicates using EZ DNA Methylation-Lightning Kit (Zymo Research, Irvine, CA) according to manufacturer's protocol. Eluted DNAs for each *FAD3* sample were mixed together and brought in equal volumes (24ul). PCR reactions were performed using 3ul mixed DNA sample for the amplification of each region of interest. A hot start Platinum Taq DNA Polymerase was used to prevent non-specific amplification (Invitrogen, Carlsbad, CA). Primer sequences are shown in Table 2.2. The parameters for the bisulfite PCR was as follows, 95°C for 5 min, followed by 5 cycles of 95°C for 1min, 51°C for 1.5min, 72°C for 2min, then 35 cycles of 95°C for 45s, 51°C for 1min, 72°C for 1.5min, followed by 72°C

for 15min, and an ending hold at 4°C. PCR products were cloned into the pGEM-T easy Vector (Promega, San Luis Obispo, CA), and 10 clones were sequenced to determine the methylation status of each region. Sequencing reactions were carried out at the DNA Core Facility (University of Missouri, Columbia, MO). Analysis of bisulfite sequencing data was performed using the online CyMATE software platform (<http://www.cymate.org/>). This experiment was repeated once.

Small RNA sequencing

T5 mid-mature soybean seeds of a transgenic plant were harvested into liquid nitrogen, and then stored at -80 for later use. Total RNA from each seed was extracted with TRIzol reagent (Invitrogen, Carlsbad, CA) and further purified with DNA-Free RNA kit (Zymo Research, Irvine, CA) to remove genomic DNA contamination. 2.5ug of RNA was submitted to the DNA Core (University of Missouri, Columbia, MO) at a concentration of 250ng/ul in nuclease-free water for library construction and small RNA sequencing. Each library was prepared and *barcoded* using TruSeq Small RNA Sample Preparation Kit (Illumina, San Diego, CA) and sequenced in the same lane of the Illumina HiSeq 2000 sequencing platform.

The resulting sequences were first trimmed off adapter sequence and filtered on length and quality. Small RNAs were mapped to the Soybean genome using Bowtie software (<http://bowtie-bio.sourceforge.net>) and sequences that did not perfectly

align were discarded. The size of each library was normalized by calculating count per million (CPM) of 18 to 25 nt genome-matching small RNA reads.

5' RACE

Mid-mature soybean seeds of a T5 transgenic plant were collected and immediately frozen in liquid nitrogen, then stored at -80 for later use. Total RNA from each seed was extracted and purified as described above. The 5' rapid amplification of cDNA ends (5'RACE) assay was performed using SMARTTM RACE cDNA Amplification kit (Clontech, Mountain View, CA). First strand cDNA was synthesized in two separate reactions using 500ng purified RNA and diluted to a final concentration of 10ng/ul in Tricine-EDTA buffer according to the manufacturer's protocol. After reverse transcription, cDNAs from same samples were pooled together and 2.5ul of the mixed cDNA were used for PCR amplification by Advantage® 2 Polymerase Mix using Universal Primer A Mix (Clontech, Mountain View, CA) and gene-specific primers with 35 cycles of 95°C for 30s, 65°C for 30, 72°C for 2min. For the *FAD3A* gene, 1/50 of the first round of PCR products were then subjected to additional 25 cycles of PCR with Nested Universal Primer A (Clontech, Mountain View, CA) and *FAD3A* gene-specific nested primer. Amplification products were separated on 1% agarose gel. Fragments with expected size were gel purified and cloned into the pGEM-T Easy Vector (Promega, San Luis Obispo, CA) for sequencing. Sequencing reactions were carried out at the DNA Core Facility (University of Missouri, Columbia, MO). Sequence alignment and were accomplished using Sequencher software (<http://www.genecodes.com/>). Primer information are listed in Table 2.3.

Statistical analysis

Comparison analysis for T₅ soybean seeds fatty acid content was done using Duncan's Multiple Range Test with $\alpha = 0.01$. Comparisons between treatment and control presented in qRT-PCR analysis were done using Independent-Samples T Test with $P = 0.01$ or 0.05 . Both statistical analysis were conducted with Statistical Package for the Social Sciences (SPSS Inc., Chicago, IL, USA).

RESULTS

Low linolenic acid phenotype of *GmFAD3* RNAi lines is inheritable

In our previous study, a hpRNA-based RNAi vector pMUFAD was designed to effectively silence the three active members of soybean *FAD3* gene family (Flores et al. 2008). To ensure efficient silencing, a 318-nt conserved region which shared 100% identity with *GmFAD3A* was selected to generate the inverted repeats (IR) of a *FAD3* target sequence of the hairpin (Flores et al. 2008). In Flores's et al. work, T₃ seeds from three homozygous *FAD3* RNAi lines were analyzed for the fatty acid phenotype of the seed oil (Flores et al. 2008). All three lines displayed a significant reduction in linolenic acid content compared to that of WT control "Jack", suggesting a potent silencing of *GmFAD3* (Table 2.4) (Flores et al., 2008).

To further investigate if the low linolenic acid phenotype of *GmFAD3* RNAi lines was heritable, T₅ seeds from the same lines were profiled for the fatty acid content of the seed oil. In order to reduce the phenotypic variation among seed from one plant, a bulk of five seeds from three individual plants were sampled for each line. The RNAi lines developed by Flores et al. (2008) demonstrated stable heritability for the low linolenic acid trait. The most dramatic reduction of linolenic acid content in T₅ seed oil was once again observed in line S-24-4D with 1.1% linolenic acid content, compared to 9% in the wild type control (Table 2.4). Consistent with previous results observed in T₃ RNAi lines, T₅ S-

24-13 and S-24-15 displayed a higher linolenic acid content than S-24-4D, at the level of 3.9% and 3.1%, respectively (Table 2.4). There was no significant difference in the levels of seed linolenic acids between T₃ and T₅ homozygous lines, suggesting that RNAi of *GmFAD3* was stably inherited. Moreover, the linoleic acid contents in all tested T₅ RNAi lines were once again significantly increased as compared with the WT control due to the block in the conversion of linoleic acid precursors to linolenic precursors in the seed ($p < 0.05$). Likewise, there was no significant changes for the palmitic, stearic, oleic acid content levels among the three RNAi lines and WT control.

Silencing levels of *FAD3A*, *FAD3B* and *FAD3C* correlate to degrees of sequence homology between inverted repeat and *GmFAD3* mRNA transcripts in the RNAi lines

The full fatty acid profiles of the three RNAi lines in this study revealed that RNAi targeting *GmFAD3* was capable of producing the reduced linolenic acid phenotype. This reduction is presumably due to the silencing of the *FAD3* gene family targeted by RNAi. To confirm this silencing at the transcript level, Flores et. al (2008) performed Northern analysis of total mRNA samples of mid-mature seeds to evaluate the silencing status in the T₀ RNAi hemizygous lines (Flores et al. 2008). As expected, almost no *FAD3* transcript was detected in five out of ten lines when probed with the 318-nt conserved sequence, indicating that all three active *FAD3* gene family members were effectively silenced (Flores et al. 2008). However, by using Northern blot analysis we were unable to distinguish the silencing efficiency for individual *FAD3* genes because transcripts of the three family

members are about the same length and share a high degree of homology (Data not shown). In order to further investigate this, real-time qRT-PCR was utilized to quantify the down-regulated transcript level of each *FAD3* gene family member because of its sensitivity to discriminate closely related sequences.

The qRT-PCR analysis was performed using total mRNA samples from a bulk of three mid-mature seed (T_5) for each of the RNAi lines S-24-4D, S-24-13 and S-24-15. Seeds of three individual plants were sampled as biological replicates and mean gene expression level was measured as a ratio compared to WT for the three *GmFAD3* family members: *FAD3A*, *FAD3B* and *FAD3C* using gene specific primers. Consistent with previous results observed by Flores et. al (2008), the transcript level of *FAD3A* and *FAD3B* were drastically decreased in all three RNAi lines, ranging from 14.6% to 39.8% and 16.5% to 32%, respectively, of WT (Figure 2.1 A-B). However, the down-regulation for *FAD3C* mRNA was much less efficient than the other two genes, with the transcript level of ~40% in S-24-4D and 80% in S-24-14 and S-24-15 of WT (Figure 2.1C). The different silencing efficacies of the three *FAD3* family members correspond to the different levels of sequence homology between the 318-bp inverted repeat (IR) used in the RNAi construct and *GmFAD3* target sequences (Figure 2.1D, 2.2). This IR is 100% identical with *FAD3A* but shares 96.5% and 84.3% sequence identity with *GmFAD3B* and *GmFAD3C*, respectively. Therefore, siRNAs generated from the 318-bp IR region contains an increased number of mismatches with the targeted *FAD3B* and *FAD3C* mRNAs, reducing transcript cleavage efficiency.

In addition to the different silencing efficacies of the three *FAD3* gene family members, the three RNAi lines displayed different silencing levels even for the same gene: S-24-4D had the most significant reduction of transcripts in all three target gene families than the other two lines (Figure 2.1 A-C). To further examine the association between the silencing efficacies and low linolenic phenotypes, we compared phenotypic data from fatty acid analysis with the target transcript levels for the three RNAi lines. As expected, the highest silencing efficiency in S-24-4D correlated with the lowest linolenic acid content of 1.1%. By contrast S-24-15 which showed a moderate silencing of *FAD3A* and *FAD3B* exhibited a higher linolenic acid level of 3.1%. And S-24-13, with the lowest silencing efficiency displayed the highest linolenic acid content of 3.6% (Figure 2.1 A-C, Table 2.4). Thus far, our data demonstrated a strong association between silencing of target *FAD3* mRNA and the reduced linolenic acid phenotype, indicating that the alteration in linolenic acid content is due to a decrease in seed expressed *FAD3* enzyme activity caused by the reduced level of *FAD3* mRNA.

Transgenes are silenced in two of the three RNAi lines

To further investigate the possible cause of different silencing levels of *FAD3* genes in the three RNAi lines, we first determined hpRNA abundance by qRT-PCR analysis using total mRNA samples from a bulk of three mid-mature seed (T_5). Mean transcript level was measured from three individual plants per RNAi line using primers amplifying the intron region. Surprisingly, the hpRNA abundance in S-24-4D was approximately 50-fold higher

than the S-24-15 and S-24-13 (Figure 2.3A). Such a substantial fold difference was unexpected because it did not correspond proportionally to the target *FAD3* mRNA level and the fatty acid phenotype. The target gene silencing level that we detected earlier was only 2-3 fold higher in S-24-4D than the remaining two lines (Figure 2.1). This unexpected result prompted us to examine the expression level of the adjacent transgene, *bar* to see if its abundance was also reduced in S-24-13 and S-14-15. The *bar* transcript level in S-24-15 was about 51% relative to that of S-24-4D (Figure 2.3B) whereas in S-24-13, the *bar* transcript almost exceeded the lowest detection limit with less than 0.1% of S-24-4D (Figure 2.3B). This low expression level of *bar* agreed with the herbicide screen phenotype of S-24-13. As previously described, all three RNAi lines contained the *bar* gene as the selectable marker, which was supposed to confer an herbicide resistant phenotype. However, T₄ seedlings from the S-24-13 event displayed an ambiguous phenotype between resistant and susceptible (Figure 2.4). These unusual findings indicated that transgenes in two of the three RNAi lines were silenced due to some unknown reason, and the silencing of the transgene possibly led to a reduced efficacy of down regulating target mRNAs in a non-linear correlative manner. The results of this experiment then prompted us to ask two questions: 1) what could cause the transgene silencing in S-24-13 and S-24-15; and 2) how hpRNA expression affected the *FAD3* target gene silencing. To address these two questions, we conducted a new series of experiments.

DNA methylation suppresses transgene expression

In plants, gene silencing could be achieved through two different pathways: transcriptional silencing maintained by DNA or histone methylation and post-transcriptional silencing in the form of mRNA cleavage or protein translational repression mediated by RNAi. To test the first possibility, we employed Bisulfite sequencing to quantify the DNA methylation level of designated genomic regions in T₅ seeds of three RNAi lines. After bisulfite treatment, unmethylated Cs in a given DNA sequence would be converted into Ts, while methylated Cs remains unchanged. The methylation level of one particular position is then calculated as the percentage of Cs presented in that position from a pool of ten sequences per line. There are three types of methylations in plants, CGN, CHG, and CHH. High level of methylation at CGN position would almost always lead to the suppression of gene expression, followed by CHG with a less significant impact. While methylation at CHH position usually do not affect transcription (Chan et al. 2005).

Four regions from the two adjacent transgenes were bisulfite sequenced to search for possible DNA methylation: one located in the Glycinin promoter region which drives the expression of hpRNA, one in the reverse 318-bp IR region of hpRNA, one in the 35S promoter region which drives the expression of *bar* gene, and one in the *bar* coding sequence (Figure 2.5A). The bisulfite sequencing results revealed different methylation patterns among the three RNAi lines.

In the glycinin promoter region, S-24-13 displayed the highest methylation level of nearly 100% at the CGN position, followed by S-24-15 with a methylation level of about 85% (Figure 2.5C). In contrast, only 11% CGNs were methylated in S-24-4D within this region. Low methylation level was detected in S-24-4D and S-24-13 at CHG and CHH

position with less than 5% methylated Cs. However, S-24-15 were found to be highly methylated at these two positions, ranging from 70% to 48%, respectively. Similar bisulfite sequencing results were obtained from the reverse 318-bp IR region (Figure 2.5C). S-24-13 was again nearly 100% methylated at CGN position but only less than 15% and 10% Cs were methylated at the CHG and CHH position, respectively. The methylation level in S-24-15 was to a less degree at the CGN position than S-24-13, but this line still displayed higher methylation level at the CHG and CHH positions than the other two lines, at 57% and 38%, respectively. S-24-4D remained unmethylated at all three positions.

Within the 35S promoter region, S-24-13 displayed highest methylation level at all three positions among the three RNAi lines, ranging from 87.5% at CGN, approximately 71% at CHG to 17% at CHH position, respectively (Figure 2.5C). In contrast, none or very limited methylated Cs were found in S-24-4D and S-24-15 at any of the three positions. Within the *bar* gene coding region, S-24-13 displayed a similar methylation level at the three positions to the 35S promoter region (Figure 2.5C). While S-24-15 was more heavily methylated at the CGN and CHG position when compared to that of 35S promoter, with methylation level at 35% and 18%, respectively. S-24-4D remained unmethylated at all three positions.

The methylation analysis conducted here revealed severe DNA methylation of transgenes in two of the three RNAi lines S-24-13 and S-24-15. When compared with transgene transcript abundance obtained through qRT-PCR analysis, these two experiments together demonstrated a strong correlation between DNA methylation and transgene transcript levels in analyzed RNAi lines. For example, highly methylated Cs in both

Glycinin promoter and IR regions of hpRNA were found in S-24-13 and S-24-15, both of which also showed ~50 times less hpRNA transcripts than that of S-24-4D (Figure 2.3A; Figure 2.5C). Likewise, S-24-13, in which almost no *bar* gene transcript was detected, was heavily methylated both at 35S promoter region and *bar* gene coding sequence, compared to the other two lines (Figure 2.3A; Figure 2.5C). Thus, we concluded that the silencing of transgenes observed in S-24-13 and S-24-15 by qRT-PCR analysis was due to DNA methylation, which might further reduced the silencing efficacy of target gene by RNAi.

No DNA methylation was detected in endogenous glycinin gene

Since the Glycinin promoter used to drive the expression of hpRNA also exists in soybean, there is a possibility that the expression of endogenous glycinin gene could also be affected through DNA methylation. To investigate this possibility, promoter and coding sequence region of endogenous glycinin gene were also bisulfite sequenced (Figure 2.5B). No considerable methylation was detected in two of the three RNAi lines compared to the wild type control, while S-24-15 was about 20% methylated at all three methylation positions in the endogenous glycinin promoter region (Figure 2.5C). However, the expression level of endogenous glycinin gene remained unchanged in all three RNAi lines compared to WT (data not shown), indicating that this amount of methylation may be tolerated by the soybean glycinin gene.

Small RNA sequencing results uncovers hpRNA processing patterns in stably transformed RNAi lines

The above results demonstrated that the hpRNA transcript level was correlated with the silencing efficacy of the target gene in analyzed RNAi lines. Given the fact that target gene silencing is mediated by siRNAs generated from the hpRNA intermediate, small RNA sequencing was performed to uncover potential differences in the species, quantity and position of transgene IR-derived siRNAs in different RNAi lines. Importantly, the discovery of these sequence and binding features of siRNAs in relation to the target gene down-regulation could provide a new insight into the siRNA-mediated cleavage mechanism, facilitating the design of the most effective siRNAs.

Overall size distribution of small RNAs

A total of 12 *bar*-coded small RNA libraries were constructed from three replications of the three RNAi lines and WT control and was subjected to high-throughput sequencing. After trimming the adaptor sequences and removal of sequences that did not map to the soybean genome or that matched to noncoding structural RNAs (rRNA, tRNA, snRNA), sequencing reads of libraries varied from 4,652,538 to 12,356,079 represented by 1,084,072 to 3,130,572 distinct sequences (Table 2.5). The average frequency of the three replications of small RNAs ranged from 18 to 25 nucleotides were plotted in Figure 2.6A-B for the three RNAi lines and WT control. When distinct sequences were compared among

these plant lines, the small RNA size distribution patterns of all libraries were nearly identical, indicating that transgene produced siRNAs only had a minor impact on small RNA size profiles. The 24-nt small RNAs were dominant in sequenced samples with an average proportion of about 71% (Figure 2.6A). This result is consistent with previous findings in model plant *Arabidopsis* that this size class is rich in sequence diversity and is the most abundant size in numerous flowering plants (Axtell 2013). For the total sequence abundance, two major peaks at 21 and 24 nucleotides was found in all libraries as a result of DCL-dependent processing (Figure 2.6B). S-24-13 exhibited a slightly higher proportion of 24-nt small RNAs at about 37% compared to the other RNAi lines and WT with an average proportion of about 31%.

Size distribution of hpRNA-produced siRNAs

The size profile of 318-nt IR produced siRNAs were further analyzed by plotting the length of siRNA sequences ranging from 18 to 25 nucleotides versus the average number of distinct sequences or the average normalized abundance from three replications (Figure 2.6C-D). Since siRNAs could be generated from both strands of the dsRNA precursor, sense and antisense siRNAs were distinguishingly presented on the plus or minus side of Y-axis. When distinct sequences were examined, size distribution patterns of the three RNAi lines were still similar; however, sizes of distinct siRNAs were distributed more evenly than that of the genome-wide analysis. Particularly, 21-nt siRNAs become the dominant species on both strands followed by 22-nt siRNAs. Another minor

peak was found at 24 nucleotide as well in all three RNAi lines (Figure 2.6C). In general, the number of distinct siRNAs found on the antisense strand was slightly less than that of the sense strand for each size class. Among the three averaged RNAi line libraries, S-24-4D exhibited the highest number of distinct siRNAs in the set regardless of size classes. Especially in case of the 21-nt siRNAs, S-24-4D displayed 287 out of 298 total distinct siRNAs that can be generated from the 318-nt IR. S-24-13 and S-24-15 shared a similar number of distinct siRNAs in all cases. Unexpectedly, a peak at 21 nucleotides was observed in WT control, although the number of reads is significantly lower than that of the RNA lines.

In case of total sequences, the abundance of each sequence in a library was normalized by calculating reads per million (CPM) of 18 to 25 nucleotides genome-matched small RNAs. The abundance of siRNAs in WT is nearly negligible compared to the three RNAi lines, indicating the origin of transgene-produced siRNAs in the analyzed RNAi lines (Figure 2.6D). Size distribution patterns of the three RNAi lines were still similar on both strands as found in distinct siRNAs and the abundance of siRNAs from the antisense strand was also slightly less than that of the sense strand for every size class. However, the distribution of siRNAs had such a strong size bias that the accumulation of 21-nt siRNAs was extremely higher than other size classes followed by 22-nt and 24-nt siRNAs. This result confirmed other researchers' findings that in plants DCL4 is normally responsible for processing exogenous hpRNA supplemented by DCL2 and DCL3 (Gascioli et al. 2005; Deleris et al. 2006; Dunoyer et al. 2007; Xie et al. 2004; Hamilton et al. 2002; Tang et al. 2003). Similar with the hpRNA expression level, the accumulation

of siRNAs from S-24-4D were about 10 to 20 times higher than the other two lines for all size classes on both strands (Figure 2.6D). This strong association indicated that siRNA accumulation is directly determined by the abundance of their hpRNA precursor. S-24-13 and S-24-15 shared very similar overall siRNA abundance, except that S-24-13 accumulated more siRNAs than S-24-15 at 21 nucleotides. In WT, 21-nt small RNAs were also the most abundant size class, although quantity of which is significantly lower than that of the RNAi lines.

siRNA distribution along 318-bp IR

In order to investigate whether a few abundant siRNAs predominate in the *FAD3* siRNA-producing locus or whether the abundance is distributed among a larger number of siRNAs, small RNAs perfectly mapped to the 318-bp IR region were plotted along the sequence versus the average of their normalized abundance from three replications (Figure 2.7A). Plus Y-axis labels represent siRNAs from the sense strand of 318-bp region, while minus Y-axis indicate siRNAs found on the opposite strand. As shown in Figure 2.7, siRNAs were not evenly distributed within the 318-nt region, instead a few prominent siRNAs exhibited high abundance. All three RNAi lines shared the same high abundant siRNA-producing regions, with 3 main peaks around 80, 145, and 275 nucleotides on the sense strand and one predominant peak within 250-300 nucleotides on the antisense strand. This plotting result implies that prominent small RNAs of high abundance could be

generated from transgene siRNA loci in addition to highly distributed low-abundance distinct siRNAs.

For most of the siRNAs, no corresponding spots with similar quantities could be found on the opposite strand. This result confirmed other researcher's findings that only one strand of the siRNAs duplex (guide strand) is selected to assemble into the active RISC, the other strand (passenger strand) is cleaved for subsequent degradation (Rand et al. 2005; Matranga et al. 2005). And the selection of guide strand is not random, one strand of the siRNA duplex is consistently more favored by the AGO protein and is used to direct the repressive regulation of complementary targets (Takeda et al. 2008; Chen et al. 2008; Khvorova et al. 2003; Schwarz et al. 2003).

Consistent with the size distribution analysis, the plotting data in Figure 2.7A also provided evidence that S-24-4D not only exhibited much higher total sequence abundance than the other two RNAi lines but also displayed more distinct siRNAs. Considering the fact that hpRNA transcript level in this line is about 50 times higher than the other two RNAi lines, it is very likely that high level of substrate hpRNA in S-24-4D increased its chance of being processed by DCL proteins, which resulted in overall higher siRNA abundance and more distinct siRNAs. Specifically, among the distinct siRNAs generated from the 318-bp IR in all three RNAi lines, about 40% were only present in S-24-4D (Figure 2.7B). However, all three lines shared 79% identity of top 1000 abundant siRNAs in each line, while most of those S-24-4D specific siRNAs only showed very low abundance (Figure 2.7C). These results implies that DCL proteins mainly processed the *FAD3* hpRNA in a similar way as in all three RNAi lines, regardless of the substrate hpRNA

expression level or genotype. Nevertheless high level of substrate did increase the quantity of products and chance of random processing. Additionally, for those siRNAs of high abundance, 21nt siRNAs were obviously the predominant size class followed by 22nt and 24nt siRNAs. This is another evidence that exogenous dsRNA intermediary is mainly recognized and processed by DCL4, DCL2 and DCL3 in plants (Gascioli et al. 2005; Deleris et al. 2006; Dunoyer et al. 2007; Xie et al. 2004; Hamilton et al. 2002; Tang et al. 2003).

The above findings provided direct evidence that high level of total sequence abundance and distinct siRNAs in S-24-4D promote its efficient silencing of *FAD3A* gene, while the majority of most abundant siRNAs shared by S-24-13 and S-24-15 also ensured their silencing of *FAD3A* referred to above.

Association of hpRNA-produced siRNAs to differential silencing efficacy of target genes in RNAi lines

To further investigate the association between hpRNA-produced siRNAs and target mRNA silencing efficacy, siRNAs from the 318-bp IR of *FAD3A* were mapped to the same region of other two *FAD3* genes, respectively (Figure 2.8). As shown in Figure 2.8, the number of distinct siRNAs mapped to *FAD3B* 318-bp region was greatly reduced compared to that of *FAD3A*, while only siRNAs around 175 nucleotides share 100% homology with the same region of *FAD3C*. Particularly, the total number of distinct antisense siRNAs,

which are triggers of target gene silencing, decreased from 1371 in *FAD3A* to 49 in *FAD3C* (Figure 2.8). Moreover, the total antisense siRNA abundance also fell from 7441.89 to 34.73 CPM, 482.92 to 2.37 CPM, 340.53 to 1.93 CPM in S-24-4D, S-24-13, S-24-15, respectively (Figure 2.8, Table 2.5). As mentioned previously, the 318-bp IR used to generate *FAD3* siRNAs is 100% identical with *GmFAD3A* but only shares 96.5% and 84.3% sequence homology with *GmFAD3B* and *GmFAD3C*, respectively (Figure 2.1D, 2.2). Therefore, siRNAs generated from the 318-bp IR region contained a considerable number of mismatches especially with *FAD3C*, which might abort their function through the failure of target binding or transcript cleavage. As a result, much less noticeable changes in mRNA level was achieved for *FAD3C* than *FAD3A* and *FAD3B* (Figure 2.1A-C). However, the silencing efficacy of *FAD3B* seemed not affected by the reduced amount of functional siRNAs, probably because it still shares relatively high identity with *FAD3A*, and the amount of functional siRNAs was sufficient to conduct an efficient silencing.

Potential transitivity of Small RNAs

In addition to the potential off-target effect that siRNA generated from 318-bp IR might cause, a second specificity problem can occur via the generation of secondary siRNA from regions outside of the sequence initially targeted by trigger-derived primary siRNA. This phenomenon, termed transitive silencing, leads to the degradation of secondary targets without sequence homology with the initial silencing inducer (Eamens et al. 2008). Although transitivity along endogenous transcripts appears to occur rarely in plants, we

should not rule out and still need to pay careful attention to the possible involvement of endogenous transcripts in signal amplification of RNAi (Bleys et al. 2006; Vaistij et al. 2002; Kościńska et al. 2005; Petersen and Albrechtsen 2005; Sanders et al. 2002; Miki et al. 2005). In order to explore this possibility, perfectly matched small RNA was mapped to the three target *FAD3* transcripts. The 318-bp siRNA target region was indicated by a black horizontal line beneath the x-axis (Figure 2.9).

As shown in Figure 2.9A-C, all three RNAi lines exhibited small RNAs outside of the original 318-bp target region on the three *FAD3* transcripts and apparently 21-nt small RNAs were dominant in all cases. Particularly, small RNAs from the antisense strand implicated their possible origin from dsRNA produced via the activity of RdRP directed by IR-derived siRNAs. Moreover, more distinct small RNAs were found on the *FAD3A* transcript than *FAD3B* and *FAD3C*, which might be a result of the greater amount of primary siRNAs targeting the transcript. However, all of these small RNAs outside of original 318-bp target region exhibited low CPM around 0.1, suggesting that even if transitivity exists, it happens at a relatively low frequency. The above finding is consistent with previous studies, that endogenous sequences are protected from transitivity by some inherent feature (Bleys et al. 2006; Vaistij et al. 2002; Kościńska et al. 2005; Petersen and Albrechtsen 2005; Sanders et al. 2002; Miki et al. 2005)

In addition to small RNAs found outside of the target sequences, perfectly matched small RNAs were detected within the rice waxy-a intron region (Figure 2.9D). In this case, low abundance small RNAs were evenly distributed on both strand in all three RNAi lines. There is still no obvious difference among the three lines in terms of total small RNA

abundance. When the distinct sequences were examined, S-24-4D exhibited greater amount of 21-nt small RNAs than other size classes, while the other two lines accumulate more 22-nt small RNAs. In another study, small RNAs corresponding to the loop portion of the hairpin transcript were also detected, and such phenomenon is due to the transitive self-silencing of the hairpin transgene (Mlotshwa et al. 2008). Thus, our results provide another example that secondary siRNAs could be generated via primary stem siRNAs targeting the hpRNA itself. However, the relative low abundance of loop secondary siRNAs implicate that such transitive self-silencing is not very effective and may not be the primary cause of the silenced transgenes in S-24-13 and S-24-15.

small RNA encoding part of the transgene

Small RNA encoding part of the Glycinin promoter and 35S promoter

Previous investigation implicated that the expression of a transgene was highly suppressed by DNA methylation. According to the literature, methylation of transgene could be induced by small RNA targeting of its promoter or 5'UTR region (Eamens et al. 2008; Huettel et al. 2007; Mourrain et al. 2007). However, siRNAs generated from the 318-bp IR should have not targeted the Glycinin promoter or 35 promoter as a result of low sequence homology. In order to further investigate this, small RNA from the 318-bp IR or from the entire library were used to search for small RNAs targeting the two promoter

regions, respectively. Small RNAs perfectly mapped to the two targets were plotted along the sequence using the same approach as described in previous sections.

Small RNA mapping result indicated that no siRNAs generated from the 318-bp region also share 100% homology with either of the two promoter sequences, indicating that methylation of the transgene was not likely mediated by *FAD3* siRNAs. However, all three RNAi lines displayed significantly higher amount of small RNAs targeting the Glycinin promoter region than WT, although none of these small RNAs were expressed at high levels (CPM < 1). These small RNAs were found on both strands of the target sequence, and no significant difference was found between the non-methylated line S-24-4D and the two methylated lines S-24-13 and S-24-15 regarding to the overall abundance of small RNAs (Figure 2.10A, Table 2.5). When the distinct sequences were examined, S-24-4D exhibited greater amounts of 21-nt and 22-nt small RNAs than other size classes, while the other two lines group were closer together and displayed significantly high level of 22-nt and 24-nt small RNAs (Figure 2.10A, Table 2.5). Since most of the identified siRNAs involved in RdDM are 24-nt long, we compared 24-nt small RNAs mapped to the antisense strand of target sequence among the three RNAi lines. The comparison did reveal a few S-24-13 and/or S-24-15 specific small RNAs; however, these small RNAs were expressed at very low level (CPM < 0.2) and only one of them were present in all three replications in S-24-13. On the other side, S-24-4D exhibited greater amount of specific 24-nt small RNAs than the other two methylated RNAi lines (Figure 2.10C). Taken together, we speculate that these 24-nt small RNAs were not likely the inducers of DNA methylation in these two RNAi lines.

Similarly, significant amount of small RNAs from sources other than 318-nt were found in the 35S promoter region in the three RNAi lines compared to WT. Among the three RNAi lines, S-24-13 exhibited highest abundance of both total and distinct small RNAs (Figure 2.10B, Table 2.5). 22-nt small RNAs were again dominant on both of the strands in S-24-13, followed by 21-nt and 24-nt (Figure 2.10B, Table 2.5). In contrast, S-24-4D and S-24-15 displayed higher level of 21-nt small RNAs than other size classes. When 24-nt small RNAs from the antisense strand were examined, 56 distinct small RNAs were found in S-24-13, while only two and none was present in S-24-4D and S-24-15, respectively (Table 2.5). Given the fact that S-24-13 was severely methylated in the 35S promoter region, such a significant difference in 24-nt antisense small RNAs among the three lines may reveal the molecular basis of this phenomenon; that is, S-24-13 specific 24-nt small RNAs were probably the inducers of DNA methylation in that line.

Small RNAs share homology with endogenous Glycinin transcript

The investigation outlined above revealed that both the Glycinin promoter and 35S promoter were targeted by non-*FAD3* small RNAs in analyzed RNAi lines. In order to further investigate whether endogenous glycinin is affected by T-DNA insertion, small RNA from the 318-bp IR or from the entire library were mapped to the Glycinin gene coding region and plotted along the sequence using the same approach as described in previous section.

Consistent with previous results, no siRNAs generated from the 318-bp region displayed 100% homology with glycinin gene coding sequence (Figure 2.11). However, all three RNAi lines and WT exhibited considerable amount of non-*FAD3* small RNAs in region, although none of these small RNAs were expressed at high levels (CPM < 1). There is no significant differences between S-24-4D, S-24-13 and WT, in terms of total small RNA abundance and distinct small RNAs, while that of S-24-15 was lower than others (Figure 2.11A, Table 2.5). Within the 5' UTR region (identical sequence with transgene), all three RNAi lines displayed greater amount of small RNAs than WT, especially the two methylated lines. Otherwise, small RNAs were evenly distributed along the template (Figure 2.11A). Sequence analysis implicated that the three RNAi lines and WT share a proportion of distinct small RNAs but also displayed considerable amount of specific ones of their own (2.11B). In addition to methylation analysis results, the above investigation also provided evidence that endogenous Glycinin gene was unaffected by the insertion of transgene even if they share the same promoter. Thus, we speculate that small RNAs found in the Glycinin promoter and 35S promoter region are probably restricted to transgenes only.

Small RNAs encoding part of the bar coding sequence

In addition to small RNAs encoding part of the two transgene promoter regions, perfectly matched small RNAs were found within the *bar* coding region. Strikingly, significantly high level of total small RNAs abundance and distinct small RNAs were found within the *bar* gene coding region (Figure 2.12). Especially in S-24-15, *bar* gene-

derived small RNA displayed total reads of 13836 CPM compared to 841 CPM of 318-bp IR derived small RNAs in the same line (Figure 2.12, Table 2.5), while the other two lines still possess about 20-fold higher 318-bp small RNAs than those encoding part of the *bar* gene. The majority of high abundant *bar* gene-derived small RNAs in S-24-15 were 21 nucleotide in length (Figure 2.12, Table 2.5). When compared with previous results that *bar* gene transcript level is 50% down-regulated in S-24-15 than that of S-24-4D, we speculate that post-translational silencing of the *bar* gene might be induced by high abundant small RNAs that share sequence homology.

5' RACE revealed the cleavage site directed by siRNAs

Previous investigation using small RNA sequencing uncovered that target mRNA silencing efficacy is correlated with 318-bp siRNA accumulation in the three RNAi lines. However, details about how siRNAs abundance could have affected target mRNA cleavage is still elusive. In order to further investigate this, 5'RACE was performed to identify cleavage sites within target *FAD3* mRNAs. The sequence differences among the three *FAD3* transcripts were utilized to develop the 5'RACE assay. A common forward primer which binds to the 5' Oligo Adapter sequence was used to amplify all targets. We were successful in designing reverse primers annealing uniquely to *FAD3A* and *FAD3C*; however, the high homology between *FAD3A* and *FAD3B* and rich A/T content within the *FAD3B* unique sequence proved to be challenging for primer design specifically amplifying *FAD3B*. Therefore, only *FAD3A* and *FAD3C* were used as templates for the identification of

cleavage sites. Since the two *FAD3* mRNAs could be cleaved at any position within the 318-bp target region, corresponding PCR product sizes for *FAD3A* and *FAD3C* are 281-599 bp and 143-461 bp, respectively. Fragments with expected size were separated on a standard agarose gel and purified for sequencing.

Figure 2.13 shows inferred cleavage sites as detected by 5' RACE, with the fraction of cloned 5' RACE PCR products terminating at that position. Similar with small RNA distribution patterns, cleavage sites on *FAD3A* mRNA were not evenly distributed along the 318-bp sequence either (Figure 2.13A). Most of the inferred cleavage sites were located within the last fifty nucleotides in all three lines, indicating that this region is more prone to be cleaved by siRNAs. A total of 4 cleavage sites were conserved among the three RNAi lines, all of which exhibited relatively higher cleavage frequency than other non-conserved positions. The two major cleavage sites detected at 297-298 and 301-302 nucleotide position together displayed more than 50% (13/23) and 40% (10/25) cleavage events in S-24-4D and S-24-13, respectively. In contrast, only one predominant cleavage site at 274-275 nucleotides was identified in S-24-15, accounting for 35% of total sequenced cleavage products. Since 5' RACE was performed on products with a range of sizes, cleavage sites were also detected in WT due to natural mRNA degradation. On the other hand, all cleavage events identified in S-24-4D are within the 318 region, while, 8 out of 28 clones sequenced in WT were located outside of the 318-bp, indicating that cleavage products detected in WT could be a result of mRNA degradation in the transcriptome. And the cleavage site at 274-275 nucleotide position shared by the three RNAi lines and WT might be the position where degradation of *FAD3A* transcript most likely occur.

Similarly, most of the cleavage sites identified on *FAD3C* mRNA were located on the second half of the 318-bp sequence for all three lines with an average of ~85% (Figure 2.13B-C). However, no predominant cleavage site was detected and only one cleavage site at nucleotide position 225-226 was found in all three RNAi lines. The inferred cleavage sites in WT were more evenly distributed on both half of the 318-bp sequence. Moreover, 6 out of 25, 13 out of 27, 7 out of 24 and 12 out of 33 sequenced clones located on the outside of the 318-bp region in S-24-4D, S-24-13, S-24-15 and WT, respectively. This result implies that cleavage of *FAD3C* mRNA was not that efficient as *FAD3A* thus the sequenced clones should contain a number of natural degradation products other than those of RNAi.

Detection of preferential cleavage by 318-bp siRNAs compared with siRNA abundance in the same region suggests the effect is due to the particularly high level of corresponding antisense siRNAs. This result echoes that previous report as “hot spot” (De Paoli et al. 2009).

Putative functional siRNAs

The above investigation using 5' RACE implicated that preferred cleavage sites on *FAD3A* mRNA are located within same region producing high abundant antisense siRNAs. In order to associate 318-bp IR-derived siRNAs with identified cleavage events, siRNAs that cover either of the two cleavage sites at 297-298 and 301-302 nucleotide were selected

for analysis. Since target mRNA cleavages is almost always directed by 21-nt antisense siRNAs, only these siRNA were counted. In addition, cleavage is more likely to happen in the center of the siRNA sequence, siRNAs exhibit the two cleavage sites within the first or last two nucleotides were discarded.

A total of 20 siRNAs were found to be the potential inducer of cleavage events occurred in either of the two positions (Table 2.6). Since the abundance of siRNAs are also important for their function, these siRNAs are ranked according their CMP in S-24-4D from high to low and the corresponding expression level in each RNAi line and WT was also listed. Putative functional siRNAs were mapped to the target *FAD3A* mRNA in Figure 2.14 with red arrows indicated the two cleavage sites. We anticipate that it is likely these siRNAs are most responsible of the cleavage events detected by 5' RACE.

DISCUSSION

This study is directed at improving RNAi technology as a tool to analyze gene function and manipulate commercial traits in soybean. To fully capitalize on the potential of RNAi, the endogenous soybean gene family *GmFAD3* was chosen as a test model gene. In our lab's previous research, a hpRNA-based RNAi vector pMUFAD was designed to effectively silence the three active members of soybean *FAD3* gene family (Flores et al. 2008). High level of silencing was achieved by transgene produced siRNAs, which led to a significant reduction of linolenic acid content in the seed oil. However, variations were detected in the down-regulated linolenic acid level between different RNAi lines (Flores et al. 2008).

The work in this study furthered our understanding of hpRNA-mediated RNAi in several ways. We demonstrated that the low linolenic phenotype achieved via hpRNA-derived siRNAs could be inherited stably at least as far as the T₅ generation and the transgene is expected to persist. To date, the stability of hpRNA-mediated RNAi were only reported in a few studies. In *Arabidopsis*, the mutant trait obtained by hpRNA transgenes is inherited stably until the T₅ generation (Stoutjesdijk et al. 2002). And the hpRNA transgene in a commercial rice RNAi line appears to have been stable for over 20 generations (Kusaba et al. 2003). Our study is the first time that stability of hpRNA-induced RNAi were reported in soybean. RNAi induced by hpRNA does not require the

generation of dsRNA mediated by RdRP, which might be the reason why this type of RNA silencing is inherited more stably than other RNAi delivery method (Béclin et al. 2002).

The association between silencing of target *FAD3* mRNA and the reduced linolenic acid phenotype indicated that the alteration of linolenic acid content is presumably due to decrease in seed expressed *FAD3* enzyme activity caused by the reduced level of *FAD3* mRNA. Among the three targets, down-regulation of *FAD3C* mRNA was much less efficient than *FAD3A* and *FAD3B*. Our small RNA sequencing data clearly showed that the reduced silencing efficacy of *FAD3C* is caused by relatively low sequence homology with the 318-bp IR. Both the total siRNA abundance and the number of distinct 318-bp IR-derived antisense siRNAs decreased dramatically in *FAD3C* when compared with *FAD3A*. That means, most of the siRNAs generated from the 318-bp IR region contained a considerable number of mismatches with the targeted *FAD3C* mRNA, which possibly resulted in the failure of binding or transcript cleavage. The silencing efficacy of *FAD3B* seemed not affected by the reduced amount of functional siRNAs, probably because it still shares relatively high identity with *FAD3A*, and the amount of functional siRNAs was sufficient to conduct an efficient silencing. The low linolenic seed phenotype observed in the three RNAi lines seemed not affected by the relatively unchanged *FAD3C* transcript level. According to the literature, *GmFAD3A* has the highest expression level of the three homologs in developing seeds (William 82, soybean growth stage R5) and has been shown to be the major contributors to seed linolenic acid levels (Bilyeu et al. 2003; Bilyeu et al. 2005; Bilyeu et al. 2006; Bilyeu et al. 2011). The other two genes are much less expressed compared to *GmFAD3A* and accordingly have a less great impact on seed linolenic acid

content (Bilyeu et al. 2003; Bilyeu et al. 2005; Bilyeu et al. 2006; Bilyeu et al. 2011). Consistent with these findings, our qRT-PCR results implicated that transcript level of *FAD3A* in mid-mature seeds (Jack, soybean growth stage R6.5) was about 3-4 times higher than *FAD3B* and *FAD3C*. Thus, the low linolenic phenotype observed in the three RNAi lines is presumably due to the high silencing level of *FAD3A*.

In addition to the different silencing efficacy observed in the three target genes, the three RNAi lines also displayed different silencing level even for the same gene. Our qRT-PCR analysis demonstrated that the relatively less efficient target gene silencing observed in S-24-13 and S-24-15 is caused by 50-fold reduced transcript level of hpRNA relative to that of S-24-4D. High level of substrate hpRNA in S-24-4D resulted in overall higher siRNA abundance and more distinct siRNAs in that line as our small RNA data have demonstrated, which further increased the silencing efficacy of target gene by RNAi. In spite of the differences in siRNA abundance, all three lines shared 79% identity of top 1000 abundant siRNAs in each line, while most of those S-24-4D specific siRNAs only showed very low abundance. These results implied that DCL proteins mainly processed the *FAD3* hpRNA in a similar way in all three RNAi lines, regardless of the substrate hpRNA expression level or genotype. The above findings provided direct evidence that high level of total sequence abundance and distinct siRNAs in S-24-4D promote its efficient silencing of *FAD3A* gene, while the majority of most abundant siRNAs shared by S-24-13 and S-24-15 also ensured their silencing of *FAD3A* referred to above.

The plotting approach used in this study further revealed that prominent small RNAs of high abundance could be generated from hairpin transcript in addition to highly

distributed low-abundance distinct siRNAs. Among all the siRNAs generated from the 318-bp IR, 21nt siRNAs were obviously the predominant size class followed by 22nt and 24nt siRNAs. This result confirmed other researchers' findings that DCL4 is normally responsible for processing exogenous hpRNA supplemented by DCL2 and DCL3 (Gascioli et al. 2005; Deleris et al. 2006; Dunoyer et al. 2007; Xie et al. 2004; Hamilton et al. 2002; Tang et al. 2003). However, our sequencing data demonstrated a significant accumulation of 22-nt 318-bp IR siRNAs, that even higher than that of 24-nt siRNAs. Such accumulation of 22-nt stem siRNAs was previously reported in Mlotshwa's et al. work, where DCL2 is required for the transitive self-silencing of the hairpin transgene in *Arabidopsis* (Mlotshwa et al. 2008). Similarly to their findings, small RNAs corresponding to the loop portion of the hairpin transcript (Rice waxy-a intron) were also detected in this study (Mlotshwa et al. 2008). In addition, detailed sequencing data revealed different loop secondary siRNA accumulation patterns in the three RNAi lines, where S-24-13 and S-24-15 exhibited predominant 22-nt siRNAs while 21-nt siRNAs dominate in S-24-4D. Thus, our results provide another example that a hairpin transgene could become a target for transitive silencing activated by self-derived secondary siRNAs. The relative low abundance of loop secondary siRNAs implicate that such transitive silencing is not very effective and may not be the primary cause of the silenced transgenes in S-24-13 and S-24-15. However, given the fact that secondary siRNAs targeting the transgene are able to induce methylation of the corresponding DNA, such possibility cannot be entirely ruled out (Vaistij et al. 2002). Thus, further exploration will be needed to investigate the origin of high abundant 22-nt stem siRNAs, the possible cause of different secondary siRNA

accumulation patterns observed in the three RNAi lines and whether secondary siRNAs directed DNA methylation on the hairpin transgene.

In addition to the transitive silencing of the transgene, small RNAs outside of the original 318-bp target region on the three *FAD3* transcripts were detected in all three RNAi lines. Especially, small RNAs from the antisense strand implicated their possible origin from RdRP-dependent dsRNA precursor. However, here 21-nt small RNAs are dominant in all cases, indicating that they might be generated via the activity of DCL4. All small RNAs outside of original 318-bp target region exhibited low CPM around 0.1, suggesting that endogenous *FAD3* transcripts are less favored substrates for RdRP-dependent production of siRNAs than the hairpin transcript. The above finding is consistent with previous studies, that transitivity along endogenous transcripts is prohibited by some inherent protective feature (Bleys et al. 2006; Vaistij et al. 2002; Kościńska et al. 2005; Petersen and Albrechtsen 2005; Sanders et al. 2002; Miki et al. 2005).

The methylation analysis conducted in this study revealed severe DNA methylation of transgene promoters in two of the three RNAi lines S-24-13 and S-24-15. When associated with qRT-PCR results, the low level of transgene transcript observed in S-24-13 and S-24-15 seemed to be induced by DNA methylation, which further reduced the silencing efficacy of the target gene by RNAi. In contrast, no considerable methylation was detected in both promoter and coding sequence region of endogenous glycinin gene. In addition, expression level of the endogenous glycinin gene remained unchanged in all three RNAi lines compared to WT (data not shown). These results implicated that the

endogenous Glycinin gene was unaffected by the insertion of the transgene even if they share the same promoter.

With the aid of deep sequencing, we discovered high amount of small RNAs targeting the two transgene promoter regions in the three RNAi line, although none of these small RNAs were expressed at high levels, while no significant differences was found between S-24-4D, S-24-13 and WT in terms of small RNA targeting glycinin gene coding sequence. Generally speaking, the methylated region in the transgene always exhibited greater amount of 22-nt and 24-nt small RNAs than other size classes, while within non-methylated region 21-nt and 22-nt small RNAs are predominant. Our observation is similar to what was reported in earlier work, that the 35S promoter and 5' coding regions of the *bar* gene were highly methylated in transgenic gentian (Mishiba et al. 2005). The authors proposed that no small RNAs encoding part of the 35S promoter sequence could be detected using Northern blot (Mishiba et al. 2005). However, in another study, the 35S promoter siRNAs were found to trigger homology-dependent transcriptional silencing of 35S promoter in some T-DNA insertion mutant lines (Mlotshwa et al. 2010). The author inferred that generation of 35S promoter siRNA is promoted by complex integration patterns of the T-DNA (Mlotshwa et al. 2010). Consistent with Mlotshwa's et al finding, our sequencing data showed that a total of 446 distinct small RNAs was detected within the 35S promoter region in the three RNAi lines, while the average total abundance of these small RNAs are less than 14 CPM, which might be the reason why no 35S promoter small RNA were detected in Mishiba's et al work (Mishiba et al. 2005). However, much effort is needed to uncover the mechanism for how the transgene promoter in S-24-4D as well as

endogenous glycinin promoter escape away from DNA methylation. Also, the origin of these promoter small RNAs is still elusive and needs further exploration.

In addition to small RNAs encoding part of the two transgene promoter regions, perfectly matched small RNAs were found within the *bar* gene coding region. Strikingly, small RNAs encoding part of the *bar* gene exhibited ~16-fold higher abundance than that of 318-bp IR-derived siRNAs in S-24-15, most of which are 21 nucleotide in length. Considering that the *bar* gene transcripts were ~50% reduced in S-24-15 compared to that of S-24-4D, we speculate post-transcriptional silencing of the *bar* gene might be induced by high abundant small RNAs that share sequence homology. Based on previous studies that spreading of RNA targeting could happen via the generation of secondary siRNA from RdRP-dependent dsRNA precursor, we speculate that such a high level of *bar*-derived small RNA accumulation might be part of transitive silencing activity (Bleys et al. 2006; Vaistij et al. 2002; Kościńska et al. 2005; Petersen and Albrechtsen 2005; Sanders et al. 2002; Miki et al. 2005). However, further exploration will be needed to investigate the origin of such high abundant *bar* gene small RNAs and whether these small RNAs induced PTGS of *bar* gene in S-24-15.

Previous investigation using small RNA sequencing uncovered that target mRNA silencing efficacy is correlated with 318-bp siRNA accumulation in the three RNAi lines. In *FAD3A*, detection of preferential cleavage sites within the high abundant siRNA accumulation region suggests such effect is due to the particularly high level of corresponding antisense siRNAs. However, 5' RACE failed to reveal preferred cleavage events within the same region where 318-bp siRNAs share 100% homology with *FAD3C*

(nucleotide 171-202). Among the three RNAi lines, only S-24-4D exhibited relatively higher cleavage frequency between nucleotide 188-189. In contrast, one cleavage site between nucleotide 225-226 is universal among the three RNAi lines. Such observations in FADC together suggest that plant siRNAs may also trigger gene silencing with limited complementation, as previously demonstrated (Jackson et al. 2006; Senthil-Kumar et al. 2007; Qiu et al. 2005). Such limited sequence specificity can potentially increase the chance of off-target effect that siRNAs may silence non-target genes (Small 2007). Consistent with other studies, in this work, degradation sites were detected in WT due to relatively broad size range of 5' RACE PCR products (Alemán et al. 2007). In order to associate 318-bp siRNAs with identified cleavage events, we've identified 20 siRNAs that cover either of the two preferred cleavage sites at nucleotide 297-298 and 301-302 on *FAD3A* transcript. Further experiment will then be needed to investigate which siRNAs are most responsible for the detected cleavage events in 5' RACE.

One specific issue we experienced in this work is that samples used for small RNA sequencing suffer from cross-contamination, as transgene derived small RNAs were also found in WT control. Such contamination may happen during preparation for the small RNA libraries, and the high sensitivity of deep sequencing was taken advantage of to reveal this problem. In a recently published article, contaminations from exogenous RNA was experimentally evaluated, that even low amount of plant miRNAs could be detected in human samples (Tosar et al. 2014). The author further analyzed sequencing data generated by other researchers, and found that contaminant sequences were ubiquitous (Tosar et al. 2014). Indeed, contamination with nucleic acids from unrelated organisms could be easily

removed, while cross-contamination between samples from the same organisms may foster concern about the validity of sequencing results. In this study, the abundance of foreign small RNAs in WT are negligible when compared to the three RNAi lines. Thus, the validity of this part of work should not have been affected.

FIGURES

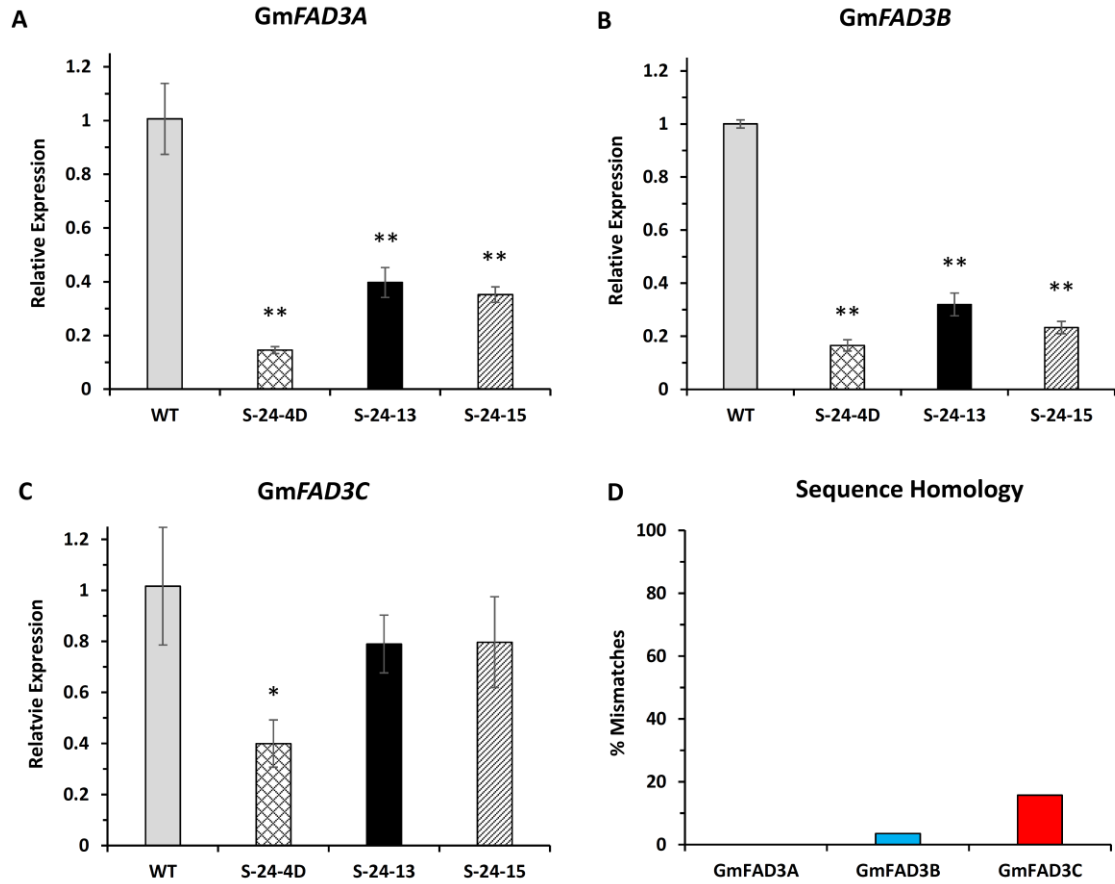


Figure 2.1 *FAD3* gene expression level of T₅ homozygous RNAi lines. Data are averages of biological triplicates \pm SD normalized to CONS7 mRNA. Independent-Samples T Test was used to test the significance. Asterisks indicate significant differences in relative expression between transgenic lines and control. (*, $p < 0.05$; **, $p < 0.01$). S24-4D, S-24-13, S-24-15 are T₅ homozygous for the pMUFAD transgene. Soybean cultivar Jack is used as wild-type control. (A) to (C) Normalized gene expression level of *GmFAD3A*, *GmFAD3B* and *GmFAD3C*, respectively. The values of wild-type plants were arbitrarily fixed to 1.0. (D) Percentage of mismatches between the 318-bp IR and corresponding regions in *GmFAD3A*, *GmFAD3B* and *GmFAD3C*, respectively.

318-nt IR	1	TGCTCAAGCTCTATGGAATCCATATTGGATATTTGTTATGTGGCTGGACTTTGTCACAT
GmFAD3A	701	TGCTCAAGCTCTATGGAATCCATATTGGATATTTGTTATGTGGCTGGACTTTGTCACAT
GmFAD3B	713	TGCTCAAGCTCTATGGAATCCATATTGGATATTTGTTATGTGGCTGGACTTTGTCACAT
GmFAD3C	704	TGCTCAAGCTCTATGGGGTTCCCTATTTGATCTTCGTCTATGTGGCTGGATTTCTGCACCT
318-nt IR	61	ACTTGCATCACCATGGTCATCATCAGAAACTGCCTTGGTATCGCGGCAAGGAATGGAGTT
GmFAD3A	761	ACTTGCATCACCATGGTCATCATCAGAAACTGCCTTGGTATCGCGGCAAGGAATGGAGTT
GmFAD3B	773	ACTTGCATCACCATGGTCACCCACCCAGAAACTGCCTTGGTACCCGCGGCAAGGAATGGAGTT
GmFAD3C	764	ACTTGCATCATCATGGTTACAAGCAGAAACTGCCTTGGTACCCGTGGCCAGGAATGGAGTT
318-nt IR	121	ATTTAAGAGGTGGTCTCACAACTGTGGATCGTGACTATGGTTGGATCAATAACATTCACC
GmFAD3A	821	ATTTAAGAGGTGGTCTCACAACTGTGGATCGTGACTATGGTTGGATCAATAACATTCACC
GmFAD3B	833	ATTTAAGAGGTGGCCCTCACCCACTGTGGATCGTGACTATGGTTGGATCAATAACATTCACC
GmFAD3C	824	ATCTAAGGGTGGTCTTACAACAGTAGATCGCGACTATGGTTGGATCAACAACATTCACC
318-nt IR	181	ATGACATTGGCACCACATGTTATTCACCATCTTTTCCCTCAAATTCCTCATTATCACCTCG
GmFAD3A	881	ATGACATTGGCACCACATGTTATTCACCATCTTTTCCCTCAAATTCCTCATTATCACCTCG
GmFAD3B	893	ATGACATTGGCACCACATGTTATTCACCATCTTTTCCCCAAATTCCTCATTATCACCTCG
GmFAD3C	884	ATGACATTGGCACCACATGTTATTCATTCACCTTTTCCCTCAAATTCCAACATTATCATTTAA
318-nt IR	241	TTGAAGCGACACAAGCAGCAAAATCAGTTCCTGGAGAGTATTACCGTGAGCCAGAAAGAT
GmFAD3A	941	TTGAAGCGACACAAGCAGCAAAATCAGTTCCTGGAGAGTATTACCGTGAGCCAGAAAGAT
GmFAD3B	953	TTGAAGCGACACAAGCAGCAAAACAGTTCCTGGAGATTACTACCGTGAGCCAGAAAGAT
GmFAD3C	944	TCGAAGCGACTAAAGCAGCAAAAGGCAGTCTAGGAAAGTATTATCGTGAGCCTCAGAAAT
318-nt IR	301	CTGCACCATTACCATTTC
GmFAD3A	1001	CTGCACCATTACCATTTC
GmFAD3B	1013	CTGCCATTACCATTTC
GmFAD3C	1004	CTGGGCCATTGCCATTTC

Figure 2.2 Alignments of nucleotide sequence of the 318-nt inverted repeat (IR) with partial cDNAs of *GmFAD3A*, *GmFAD3B*, and *GmFAD3C* (genotype “Jack”). Conserved nucleotides are highlighted in black. The position of each *FAD3* sequence is marked beginning from start codon and indicated at the beginning of each line of the alignment. Sequence alignment were accomplished using CLUSTALW2 (<https://www.ebi.ac.uk/Tools/msa/clustalw2/>) (Flores et al., 2008).

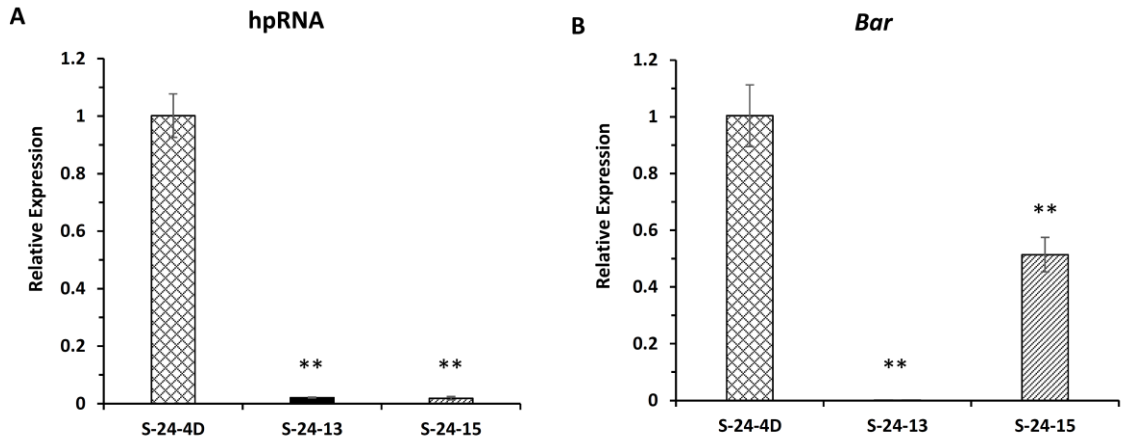


Figure 2.3 hpRNA and Bar transcript level of pMUFAD homozygous lines. Data are averages of biological triplicates \pm SD normalized to CONS7 mRNA. Independent-Samples T Test was used to test the significance. Asterisks indicate significant differences in relative expression ($*p < 0.05$, $**p < 0.01$). S-24-4D, S-24-13, S-24-15 are T₅ homozygous for the pMUFAD transgene. Soybean cultivar Jack is used as wild-type control. **(A)** and **(B)** Normalized transcript level of hpRNA and bar, respectively. The values of S-24-4D plants were arbitrarily fixed to 1.0.

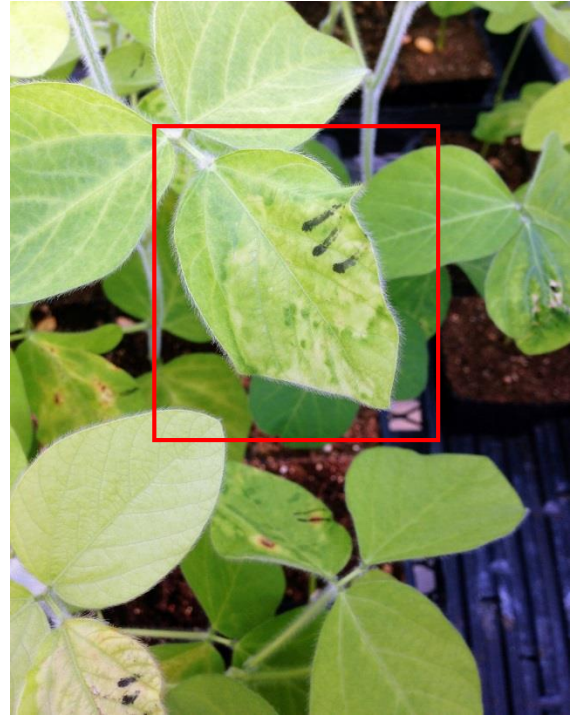
A**B**

Figure 2.4 Leaf painting phenotype comparison on seedling of T₄ S-24-13 and WT. Painted leaves were highlighted with red rectangle. Pictures were taken three days after painting. Herbicide glufosinate was applied at a concentration of 100mg/L. (A) WT; (B) S-24-13. Seedling of T₄ S-24-13 displayed an ambiguous phenotype between resistant and susceptible.

(Figure 2.5 continued)

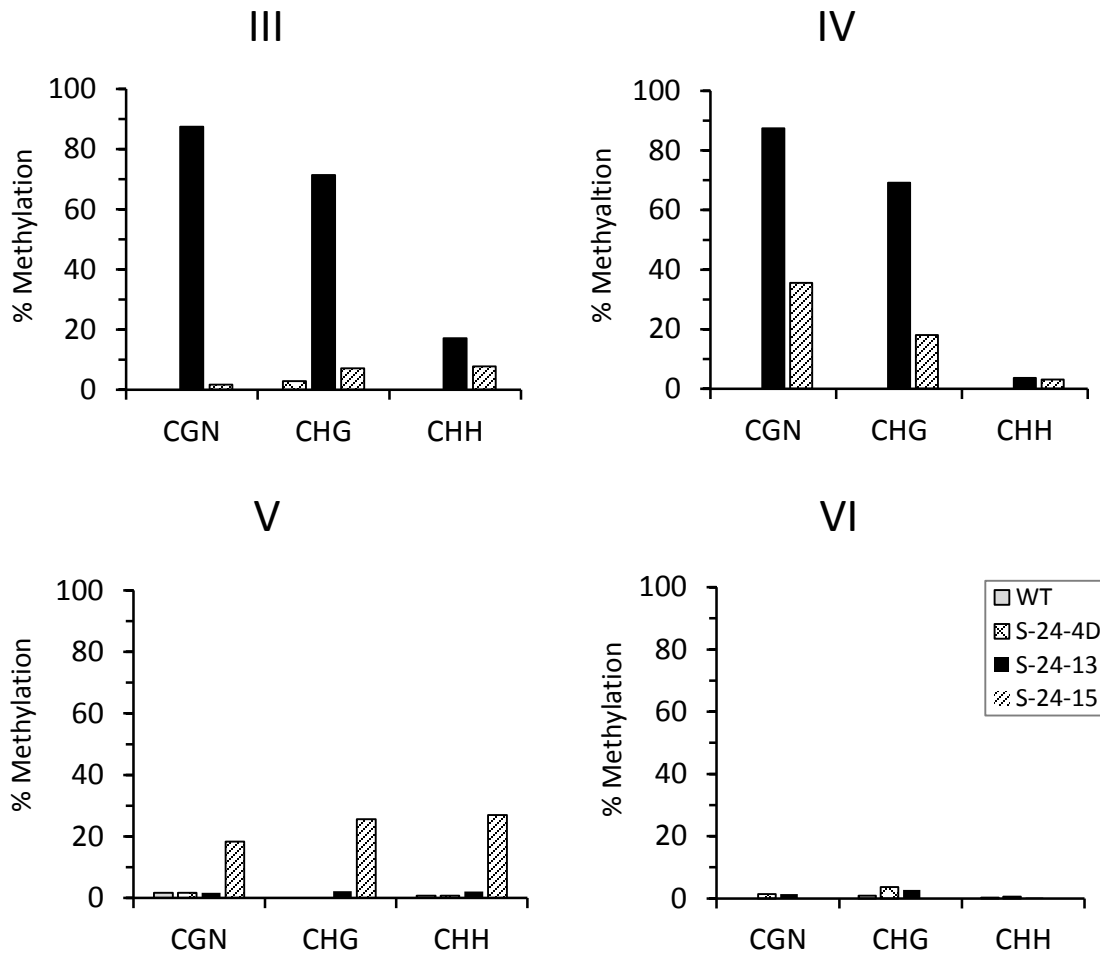


Figure 2.5 DNA methylation analysis of T₅ homozygous RNAi lines. **(A)** Schematic presentation of the T-DNA region of the plant transformation vector, pMUFAD (Flores et al., 2008). The expression cassette for the RNAi of *GmFAD3* is highlighted in grey. LB and RB represent T-DNA left and right borders, respectively; Tvsp represent soybean vegetative storage protein gene terminator; *bar* indicate bialaphos resistance gene; TEV indicate tobacco etch virus translational enhancer; CaMV35S stand for Cavliflower mosaic virus 35S promoter; OCS 3' stand for octopine synthase gene terminator; IR-R and IR-F represent the 318-bp inverted repeats of *GmFAD3* target sequence in reverse and forward directions, respectively; Rice Waxy-a Intron indicate rice Waxy-a gene intron; GlyP stand for soybean glycinin gene promoter (Flores et al., 2008).

(Figure 2.5 continued)

Lines beneath the schematic represent region I, II, III and IV examined by bisulfite sequencing. **(B)** Schematic of the *Glycinin* gene (GenBank: AB113349.1). The darkened rectangle represents exon and horizontal line represents intron. The black arrow indicates transcription starting site. Lines beneath the schematic represent region V and VI examined by bisulfite sequencing, with numbers indicating the corresponding position. **(C)** Methylation status of a 338 bp region I (294bp Glycinin promoter, 32bp vector backbone, 12bp forward inverted-repeat of the *FAD3* hair-pin), a 276bp region II, a 281bp region III and a 352bp region IV within the plant transformation vector pMUFAD; a 297 bp region V (-295 bp to 2bp) and a 255 bp region VI (1105 bp to 1359bp) within soybean *Glycinin* gene (GenBank: AB113349.1). Bar heights represent the percentage of methylation at each CGN, CHG and CHH (where N= A, T, G or C; H= A, T, or C) cytosines of 10 clones analyzed by bisulfite sequencing. Two biological replications were performed and similar results were obtained.

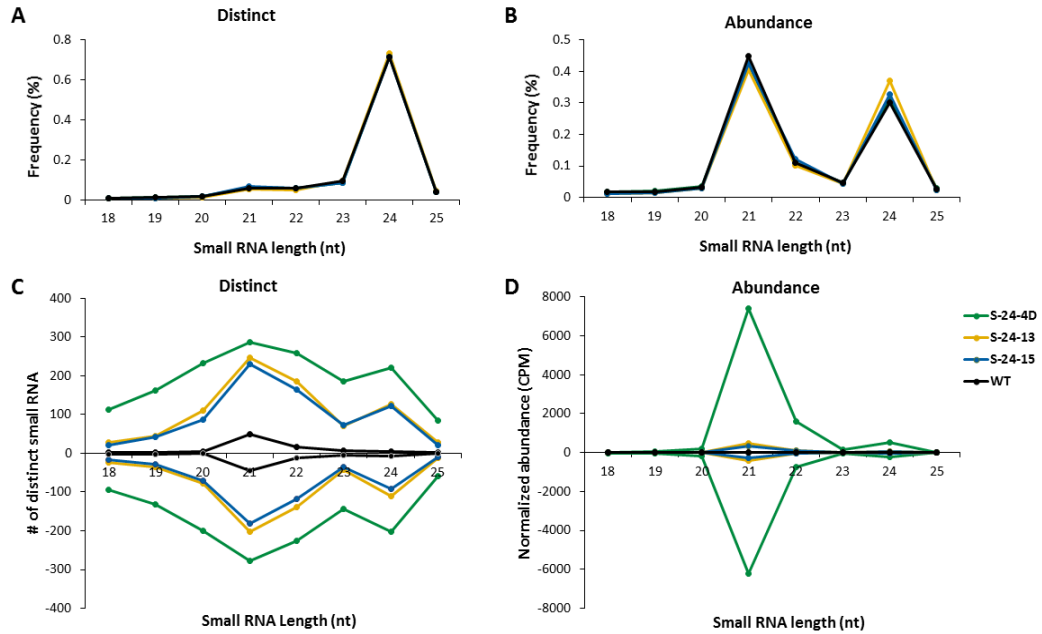


Figure 2.6 Small RNA size distribution. The size of small RNA from each sample was plotted versus frequency among distinct sequences (**A**) or total sequences (**B**) to eliminate the bias of different sequencing depth. (**C**) and (**D**) Size profiles of the 318-bp IR small RNAs for distinct sequences and total sequences, respectively. Data are average of three replication.

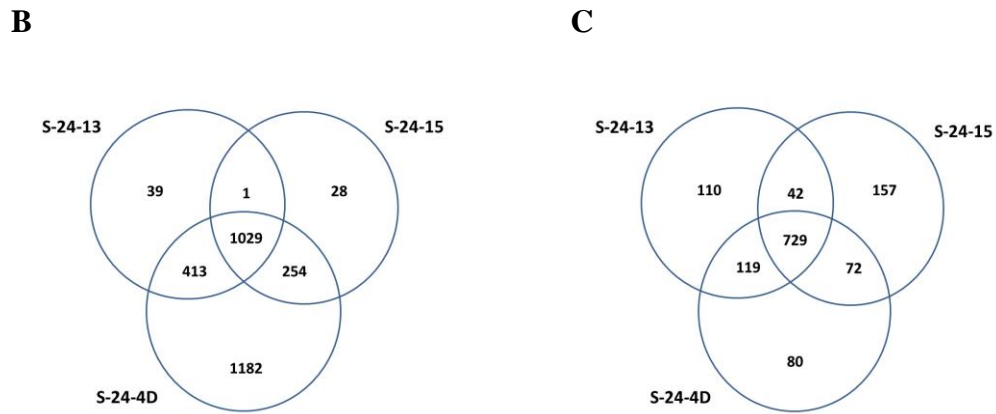
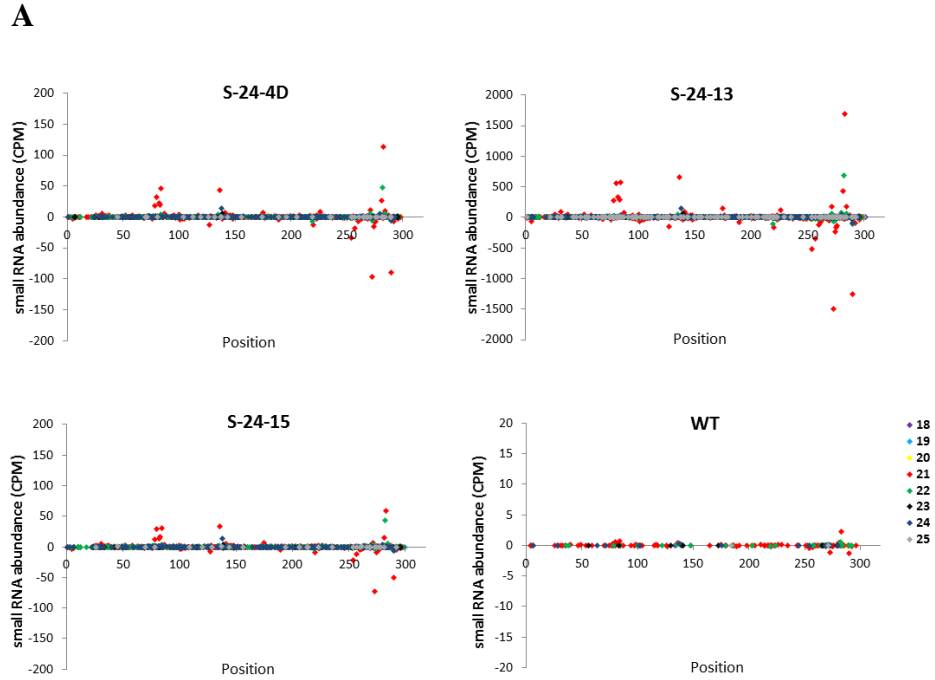
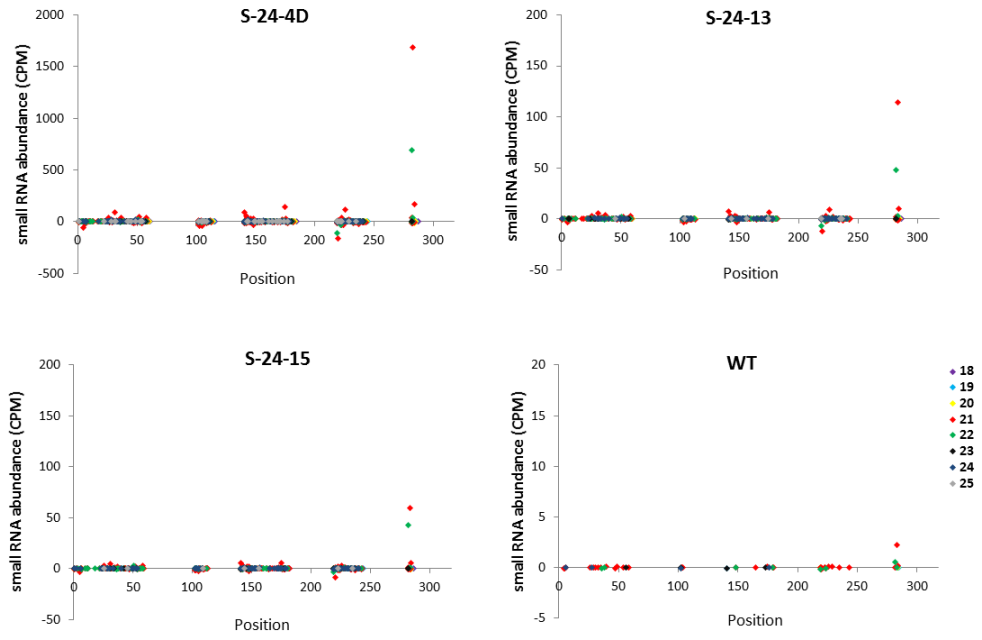
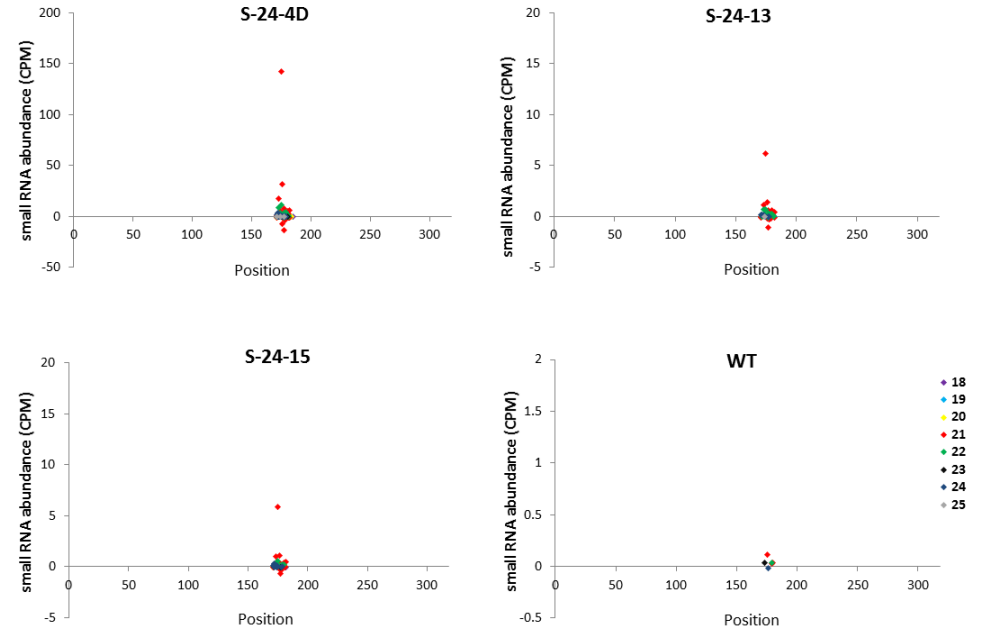


Figure 2.7 Comparison of 318-bp IR-derived siRNAs in the three RNAi lines. **(A)** Small RNAs matching the 318-bp IR were plotted versus the average of their normalized abundance from three replications. For visual clarity, the Y-axis of each diagram is adjusted based on the corresponding small RNA abundance. **(B)** and **(C)** Venn diagram represents common and specific reads from total and top 1000 abundant small RNAs in S-24-4D, S-24-13 and S-24-15, respectively.

A



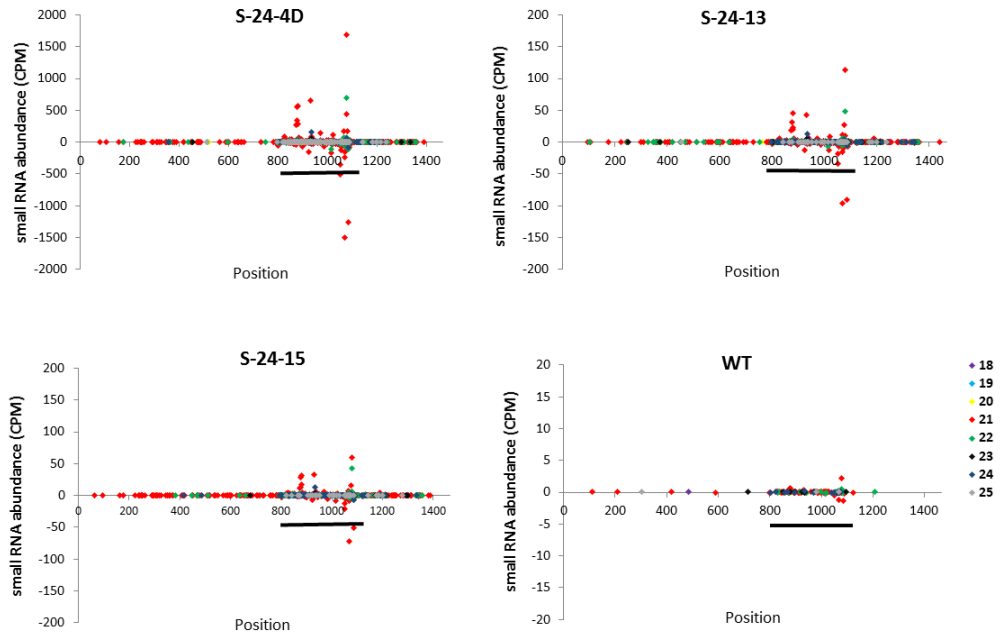
B



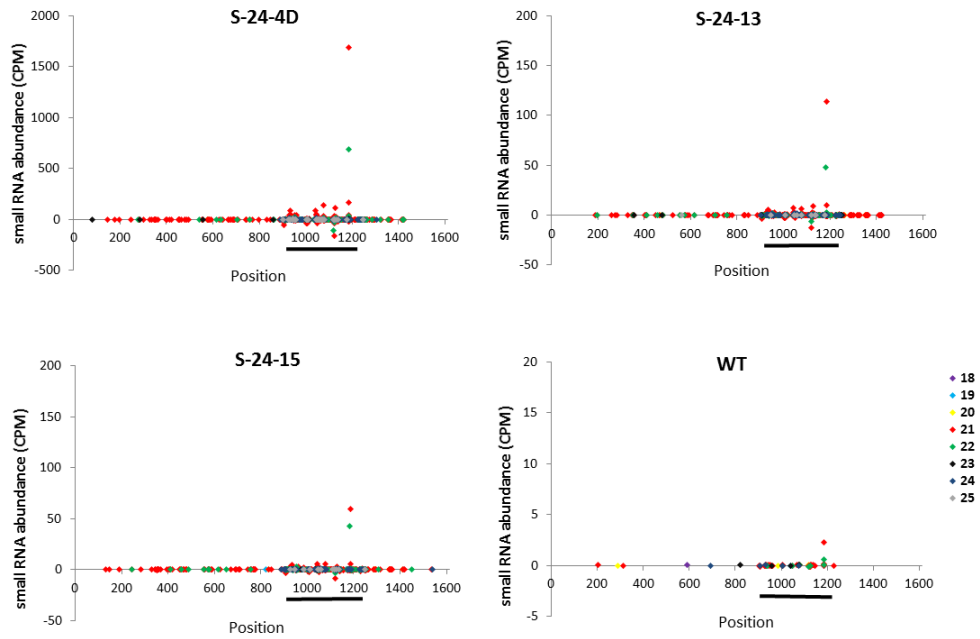
(Figure 2.8 continued)

Figure 2.8 318-bp IR-derived siRNAs targeting *FAD3B* and *FAD3C*. **(A)** and **(B)** Small RNAs generated from the 318-bp IR matching the corresponding *FAD3B* and *FAD3C* target regions were plotted versus the average of their normalized abundance from three replications, respectively. For visual clarity, Y-axis of each diagram is adjusted according to the small RNA abundance.

A

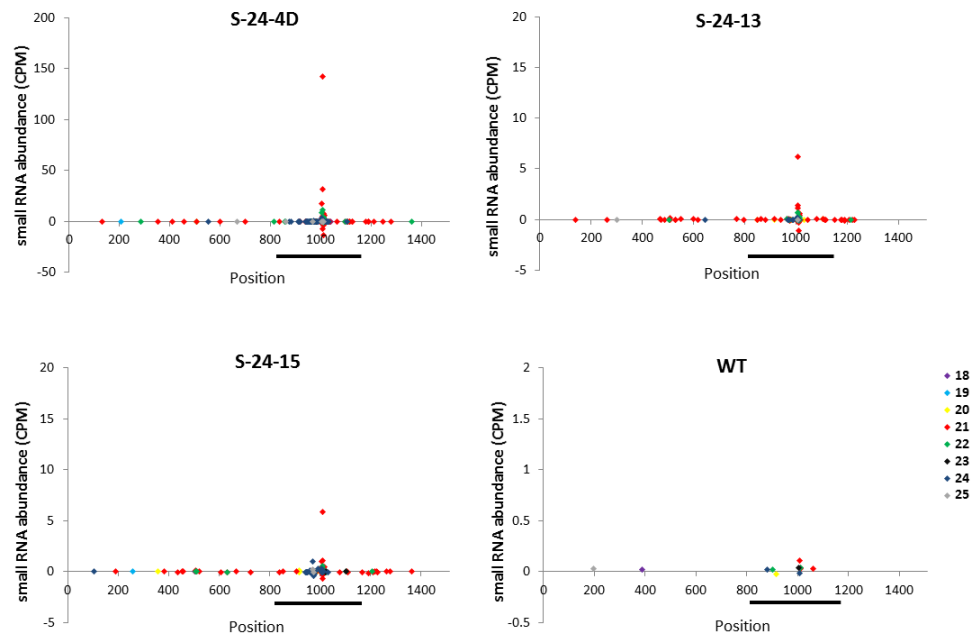


B

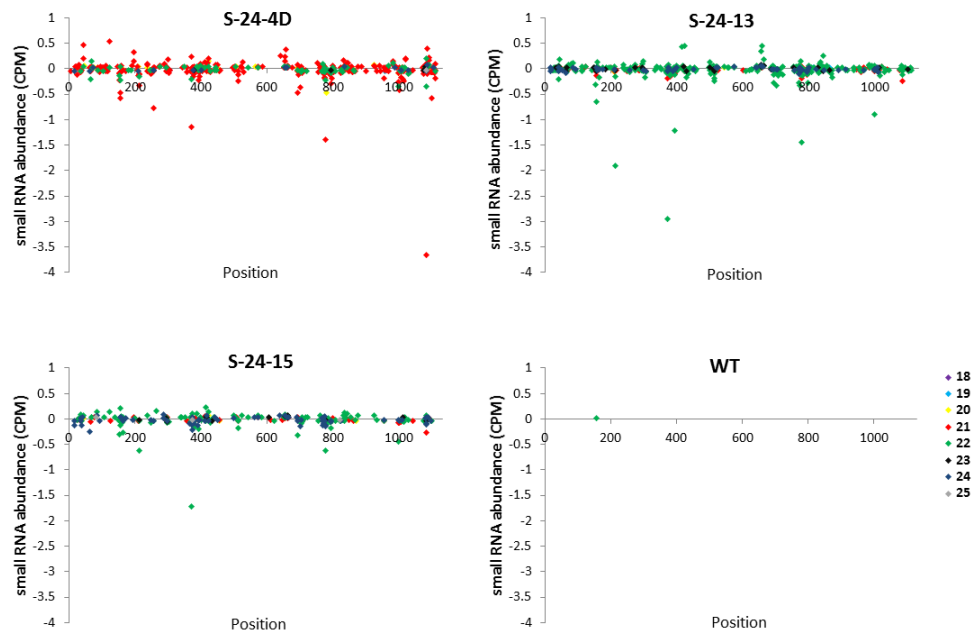


(Figure 2.9 continued)

C



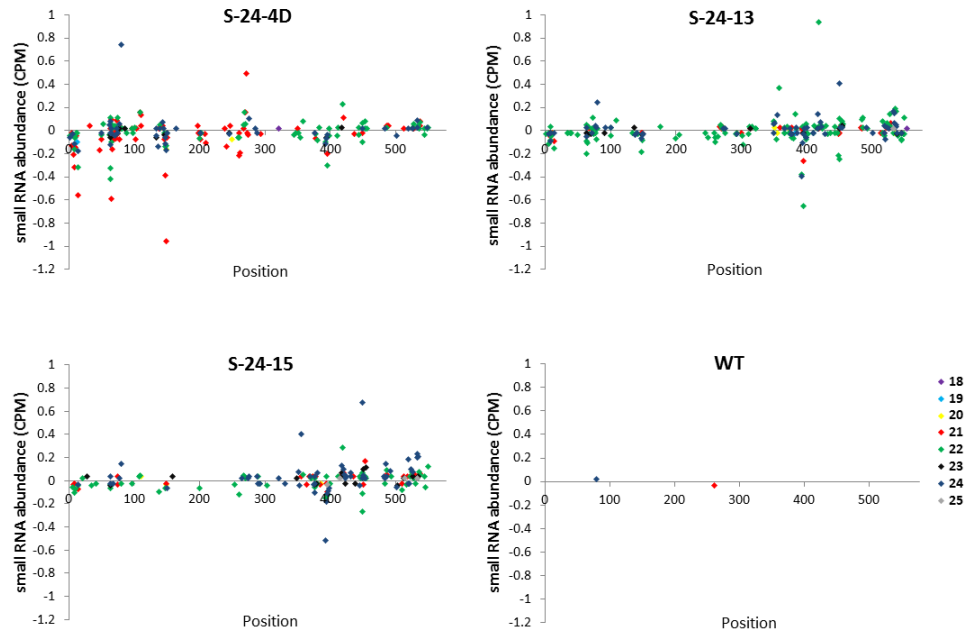
D



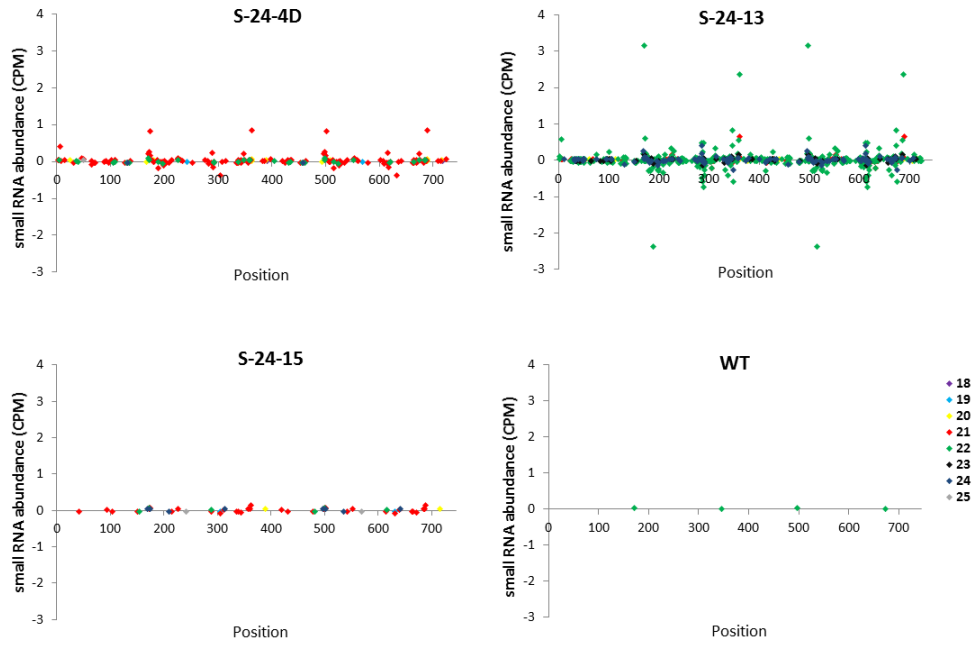
(Figure 2.9 continued)

Figure 2.9 RNAi-induced transitivity. **(A) - (D)** Small RNAs matching the *GmFAD3A*, *GmFAD3B*, *GmFAD3C* transcript sequences and the rice waxy-a intron were plotted versus the average of their normalized abundance from three replications, respectively. The 318-bp siRNA generating IR on *GmFAD3A* and corresponding target regions on *GmFAD3B* and *GmFAD3C* are indicated with black lines. For visual clarity, Y-axis of each diagram is adjusted according to the small RNA abundance.

A



B



(Figure 2.10 continued)

C

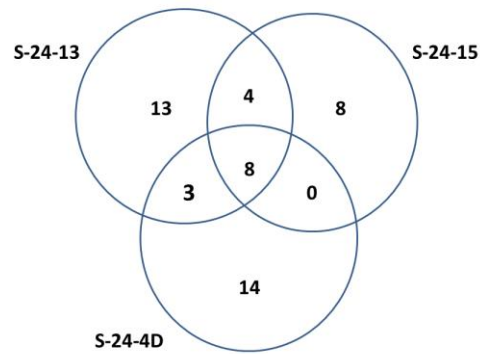


Figure 2.10 small RNA encoding part of transgene promoters. Small RNAs matching Glycinin (**A**) and 35S (**B**) promoter regions were plotted versus the average of their normalized abundance from three replications, respectively. (**C**) Venn diagram represents common and specific reads from 24-nt antisense small RNAs matching Glycinin promoter region in S-24-4D, S-24-13 and S-24-15, respectively.

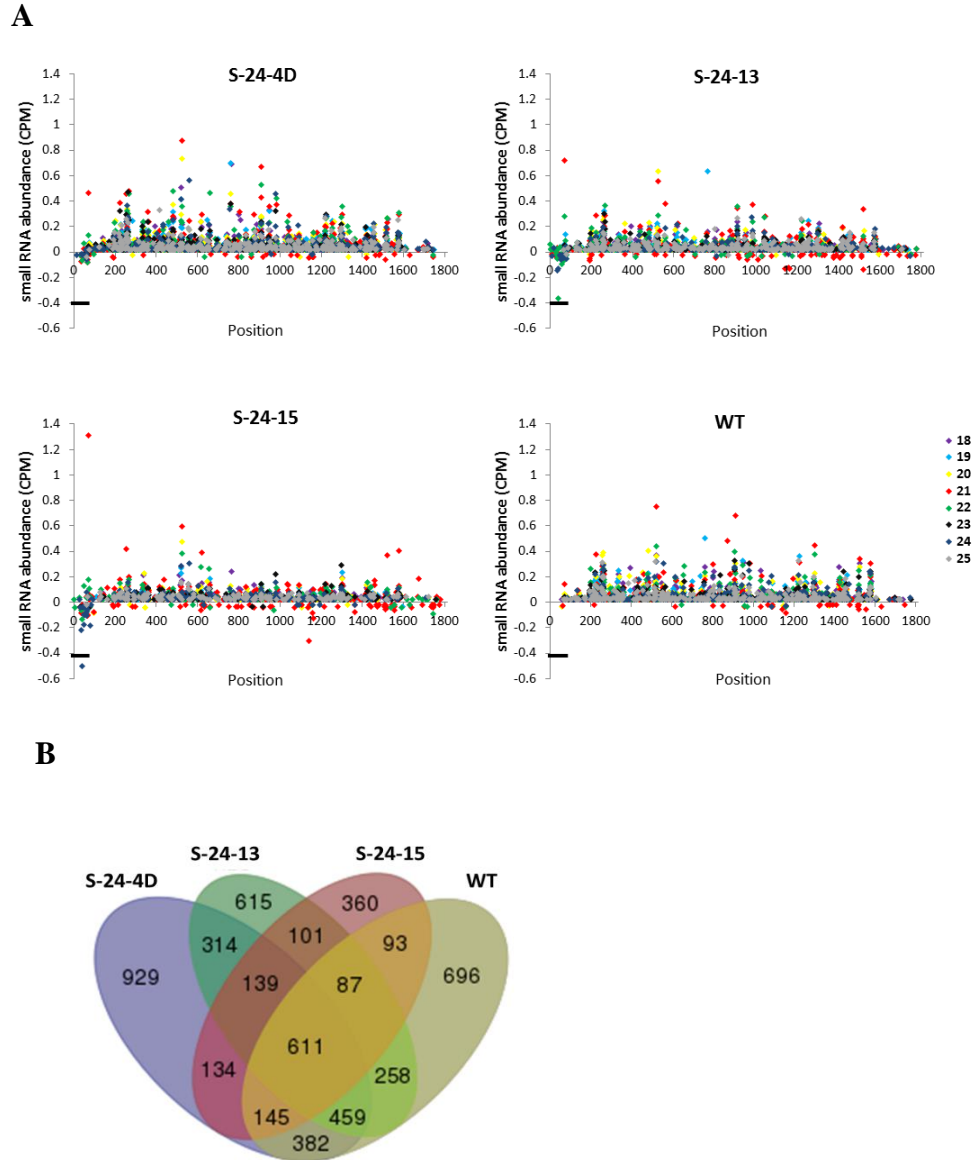


Figure 2.11 Small RNAs matching endogenous *Glycinin* transcript. (A) Small RNAs matching endogenous *Glycinin* gene coding sequence were plotted versus the average of their normalized abundance from three replications, respectively. The 5' UTR region (identical sequence with transgene) was indicated with black lines. (B) Venn diagram represents common and specific reads from small RNAs matching endogenous *Glycinin* gene coding sequence in S-24-4D, S-24-13 and S-24-15, respectively.

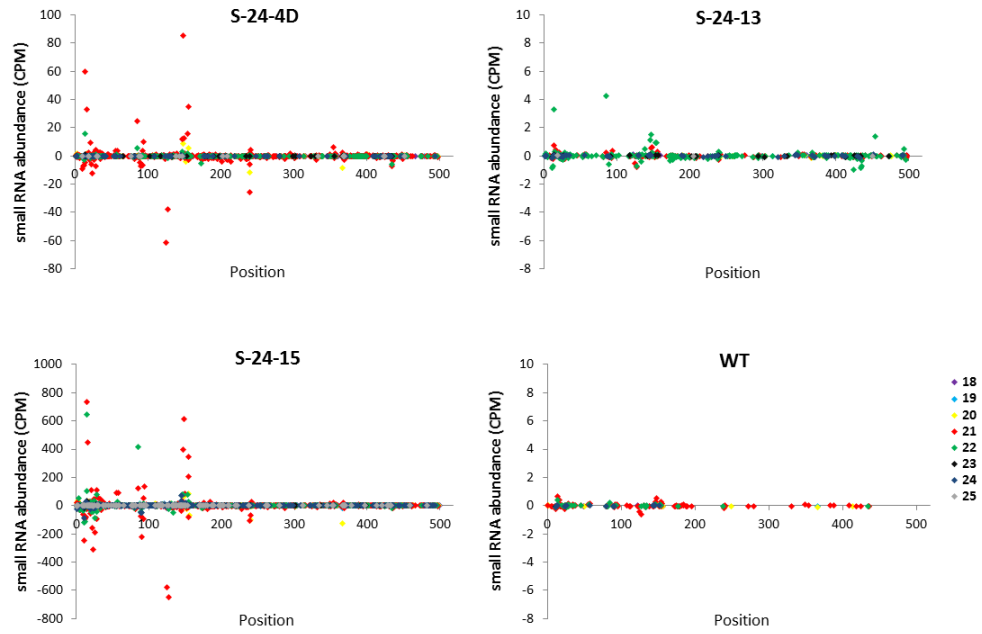
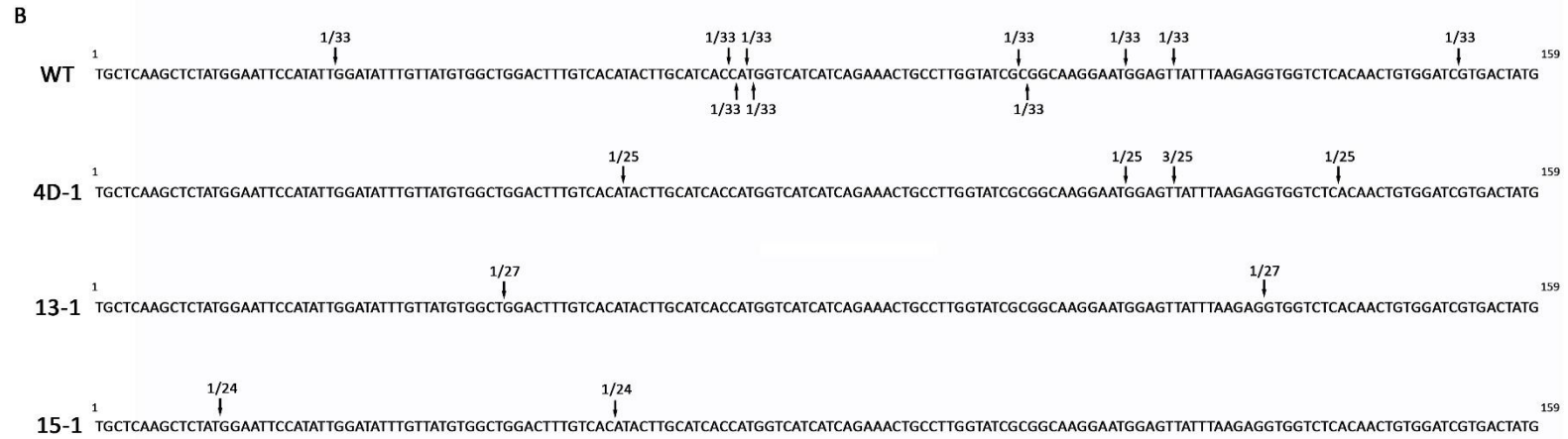


Figure 2.12 Small RNAs encoding part of the *bar* coding sequence. Small RNAs matching the *bar* gene coding sequence of pMUFAD were plotted versus the average of their normalized abundance from three replications, respectively. For visual clarity, Y-axis of each diagram is adjusted according to the small RNA abundance.

(Figure 2.13 continued)



(Figure 2.13 continued)

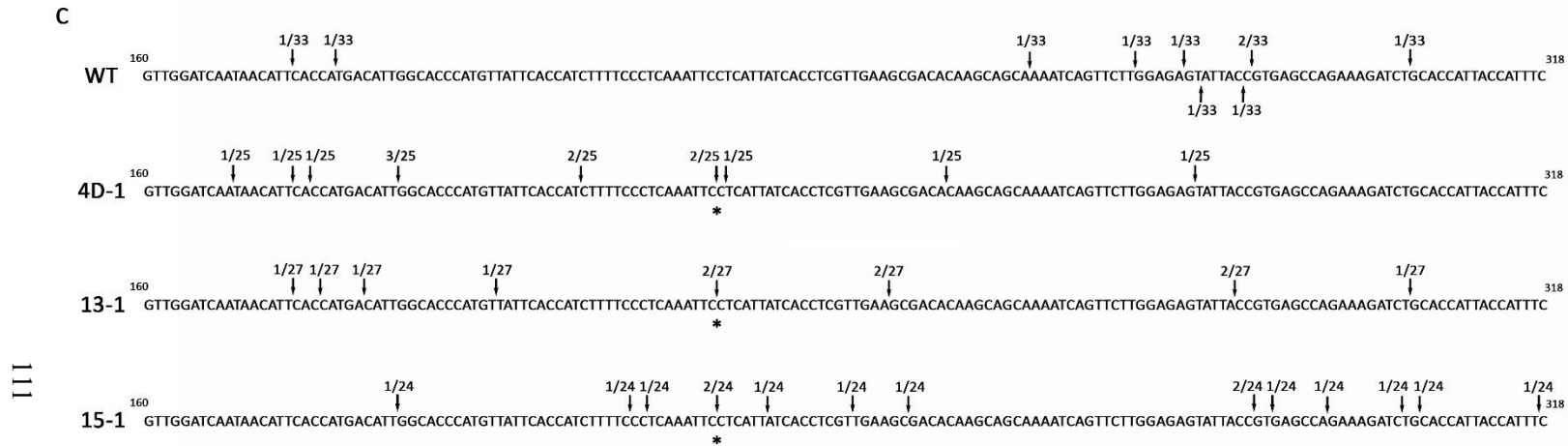


Figure 2.13 5'RACE on *GmFAD3A* and *GmFAD3C* mRNAs in T_5 RNAi lines. Arrows indicate the inferred cleavage sites and numbers above represent the fractions of cloned 5' RACE PCR products terminating at this position. Degradation sites detected with high frequency are highlighted in red, and those present across the three RNAi lines are highlighted with asterisks. (A) Summary of the 5' RACE analysis performed on the 318-bp region of *GmFAD3A* mRNA. (B) and (C) Summary of the 5' RACE analysis performed on the corresponding region of *GmFAD3C* mRNA.



Figure 2.14 Putative functional siRNAs. 21-nt antisense 318-bp IR-derived siRNA sequences that cover either of the two cleavage sites at 297-298 and 301-302 nucleotide were mapped to the 318-bp template. Rank order is based on the corresponding CPM from high to low in S-24-4D. siRNAs cover the two cleavage sites within the first or last two nucleotides were not counted.

TABLES

Table 2.1 Sequences of primers used for qRT-PCR analysis.

Gene	Primer name	Primer sequence
<i>FAD3A</i>	qRT- <i>FAD3A</i> -F	AGCGACACAAGCAGCAAAAT
	qRT- <i>FAD3A</i> -R	GTCTCGGTGCGAGTGAAGGT
<i>FAD3B</i>	qRT- <i>FAD3B</i> -F	CCCACCCAGTGAGAGAAAA
	qRT- <i>FAD3B</i> -R	AGCACTAGAAGTGGACTAGTTATGAAT
<i>FAD3C</i>	qRT- <i>FAD3C</i> -F	CTCAGAAATCTGGGCCATTG
	qRT- <i>FAD3C</i> -R	TCGCTAACGAAGTGATCCTGA
CONS7	qRT-CONS7-F	ATGAATGACGGTTCCCATGTA
	qRT-CONS7-R	GGCATTAAAGGCAGCTCACTCT
Rice waxy-a Intron	qRT-Intron-F	GCTCAAAGCTCTGTGCATCTCC
	qRT-Intron-R	CAGTTTCTTGGGTGGCTAGGG
Bar	qRT-Bar-F	GGTGGGTGTAGAGCGTGGA
	qRT-Bar-R	GCTATCCCTGGCTCGTCG

Table 2.2 Sequences of primers used in Bisulfite sequencing.

Amplified Fragment	Primer name	Primer sequence
Glycinin Promoter	meth-GlyP-F	GAAGAAAAGAAATGAAATTATGTATG
	meth-GlyP-R	AACAAAAAAAAAAAACTAACTTAACC
Glycinin CDS	meth-GlyGene-F	TAACTTCTCATCCTCTTCTTCTTC
	meth-GlyGene-R	GGAATTAAGAGTAAGAGTTTTTAAAATAT
pMUFAD-Glycinin Promoter	meth-GlyP-F	GAAGAAAAGAAATGAAATTATGTATG
	meth-pMUFAD-GlyP-R	AACAAATATCCAATATAAAATTCCAT
pMUFAD-Inverted Repeat	meth-pMUFAD-318RVS-F	TAGGGGAAATGGTAATGGTG
	meth-pMUFAD-318RVS-R	CTCAAACCTCTATAAAATTCCATATT
pMUFAD-35S Promoter	meth-pMUFAD-35S-F	TTTTTTTATATAGAGGAAGGGTTT
	meth-pMUFAD-35S-R	CAATTAAAACTTTTCAACAAAAA
pMUFAD-Bar	meth-pMUFAD-Bar-F	GGATTTYAGYAGGTGGGTGTAGAG
	meth-pMUFAD-Bar-R	CATTCTTTTAAAACAAAAACAATTTT

Table 2.3 Sequences of gene specific primers used in 5' RACE.

Gene	Primer name	Primer sequence
<i>FAD3A</i>	RACE- <i>FAD3A</i>	TGTGAATGCTCTGTGCAAGTGGTAG
	RACE- <i>FAD3A</i> -NEST	TTCCATTGAGGCCCACTATGAATTCC
<i>FAD3C</i>	RACE- <i>FAD3C</i>	TAGTTGGACTGGGTCCAAGAATCTTTG

Table 2.4 Fatty acid analysis of T3 and T5 soybeans for homozygous RNAi lines.

Line name	Fatty acid content (%)				
	16:00	18:00	18:01	18:02	18:03
T ₃ ¹					
WT	10.2	3.4	16.6	58.9a ⁴	10.6a
S-24-4D	10.9	3.7	16.1	68.1b	1.2c
S-24-13	9.9	3.7	18.2	65.8b	2.4c
S-24-15	10.2	3.6	15.3	67.4b	3.6b
T ₅ ²					
WT	10.6 ± 0.6 ³	3.7 ± 0.2	18.2 ± 3.1	58.4 ± 1.8A ⁵	9.0 ± 0.7A
S-24-4D	11.0 ± 0.3	4.0 ± 0.8	17.2 ± 2.1	66.7 ± 1.7C	1.1 ± 0.1B
S-24-13	10.4 ± 0.3	3.9 ± 0.5	18.9 ± 1.6	63.0 ± 1.4B	3.9 ± 0.2C
S-24-15	10.2 ± 0.4	3.7 ± 0.4	22.4 ± 2.5	60.7 ± 2.7AB	3.1 ± 0.2D

¹ Fatty acid profile data for T₃ soybeans were obtained from our lab's previously published paper (Flores et al., 2008). Mean fatty acid content is based on 2 to 4 replications (seeds) for each soybean line. Jack is the control line with normal fatty acid content.

² Fatty acid profile data for T₅ offspring seeds from the same T₃ RNAi lines were analyzed.

³ Mean fatty acid content plus and minus one standard division. The mean value and standard division are based on 3 replications of 5 seeds bulk samples for each soybean line. Mean fatty acid content is based on 5 seeds bulk samples from 3 individual plants for each soybean line. Jack is the control line with normal fatty acid content.

⁴ Means within the same column followed by the same letter were not significantly different from each other at a = 0.01 level as detected by Duncan's Multiple Range Test.

⁵ Means within the same column followed by the same letter were not significantly different from each other at P = 0.01 level as detected by Independent-Samples T Test.

2.5 Summary of distinct small RNA and total small RNA abundance for each RNA template.

		Sense				Antisense				
		S-24-4D	S-24-13	S-24-15	WT	S-24-4D	S-24-13	S-24-15	WT	
118	<i>FAD3A-318</i>	18	112 ¹ /23.82 ²	29/1.64	22/1.33	2/0.04	94/45.71	23/2.2	16/1.92	2/0.08
		19	162/62.09	44/3.94	42/2.9	3/0.08	132/31.45	36/2.33	29/1.89	2/0.07
		20	233/177.26	109/11.77	86/10.14	4/0.15	200/158.75	79/9.63	70/8.14	0/0
		21	287/7407.27	246/488.77	230/328.7	48/6.78	277/6200.15	202/402.24	182/274.39	44/5.14
		22	259/1611.58	186/123.03	164/105.43	17/1.15	227/735.06	140/45.3	119/33.98	12/0.58
		23	185/127.65	70/11.43	73/9.88	7/0.23	144/50.42	43/3.11	35/2.98	4/0.13
		24	221/539.95	127/44.02	122/40.51	5/0.39	202/217.28	110/17.55	93/16.67	8/0.25
		25	84/20.07	27/1.58	20/1.79	2/0.05	59/3.08	11/0.56	9/0.57	0/0
		Total	1543/9969.68	838/686.17	759/500.69	88/8.87	1335/7441.89	644/482.92	553/340.53	72/6.26
		<i>FAD3B-318</i>	18	55/8.41	11/0.71	11/0.41	0/0	37/3	6/0.18	4/0.17
19	70/34.24		19/2.31	17/1.17	2/0.05	53/6.66	17/0.68	8/0.32	1/0.04	
20	107/78.7		49/5.21	41/4.61	2/0.09	88/42.67	28/2.85	24/2.06	0/0	
21	134/3091.2		107/204.91	99/128.21	22/3.22	126/754.77	84/50.27	71/36.41	12/0.53	
22	113/883.27		72/63.81	65/56.77	4/0.67	92/211.09	45/13.45	43/8.66	4/0.23	
23	73/17.07		18/0.96	22/1.12	3/0.07	51/12.19	16/0.84	12/0.89	1/0.04	
24	82/99.59		49/8.5	48/6.67	1/0.04	69/27.51	28/2.58	26/2.43	3/0.07	
25	24/5.47		6/0.44	7/0.41	0/0	17/0.62	2/0.05	0/0	0/0	
Total	658/4217.94		331/286.85	310/199.39	34/4.14	533/1058.5	226/70.89	188/50.92	21/0.9	

(Table 2.5 continued)

119	<i>FAD3C-318</i>	18	6/0.25	0/0	0/0	0/0	3/0.89	0/0	1/0.08	0/0
		19	6/0.51	0/0	2/0.04	0/0	4/0.16	1/0.02	1/0.02	0/0
		20	11/2.73	5/0.24	4/0.25	0/0	9/2.02	2/0.09	3/0.15	0/0
		21	12/226.58	12/10.99	11/9.92	3/0.17	10/28.19	8/2.02	6/1.43	0/0
		22	11/35.17	9/2.77	9/1.64	1/0.04	7/2.14	3/0.15	1/0.08	0/0
		23	8/0.95	1/0.07	3/0.08	1/0.04	2/0.06	0/0	2/0.07	0/0
		24	9/7.82	5/0.5	6/0.77	0/0	8/1.19	2/0.05	3/0.1	1/0.02
		25	1/0.03	0/0	0/0	0/0	2/0.08	1/0.03	0/0	0/0
		Total	64/274.05	32/14.57	35/12.7	5/0.24	45/34.73	17/2.37	17/1.93	1/0.02
<i>GmFAD3A</i>	18	115/23.9	31/1.68	24/1.37	3/0.07	98/45.84	23/2.2	18/1.96	2/0.08	
	19	168/62.24	50/4.2	47/3.04	3/0.08	136/31.57	40/2.43	32/1.99	2/0.07	
	20	262/178.55	131/12.48	102/10.92	4/0.15	214/159.31	97/10.14	84/8.79	0/0	
	21	395/7428.45	355/505.29	343/345.44	51/6.85	388/6212.22	300/411.18	294/284.27	46/5.18	
	22	293/1615.02	220/125.41	193/107.53	18/1.17	255/738.38	167/46.71	135/34.85	12/0.58	
	23	193/128.2	75/11.66	79/10.12	9/0.28	149/50.54	49/3.27	38/3.07	4/0.13	
	24	236/540.76	138/44.57	134/41.04	5/0.39	213/217.66	118/17.91	97/16.83	8/0.25	
	25	90/20.36	31/1.7	24/1.94	3/0.07	62/3.15	12/0.58	11/0.65	0/0	
	Total	1752/9997.47	1031/706.99	946/521.41	96/9.07	1515/7458.66	806/494.42	709/352.4	74/6.3	

(Table 2.5 continued)

	18	57/8.47	11/0.71	13/0.46	1/0.04	40/3.1	7/0.19	7/0.25	0/0
	19	73/34.32	20/2.34	23/1.33	2/0.05	57/6.77	18/0.71	10/0.36	1/0.04
	20	121/79.4	57/5.39	51/4.98	3/0.11	100/43.08	39/3.16	34/2.49	1/0.02
	21	228/3099.52	192/210.77	180/135.97	24/3.26	218/765.7	153/57.41	155/45.5	13/0.54
<i>GmFAD3B</i>	22	147/884.87	92/64.61	92/57.78	4/0.67	117/214.34	66/15.09	57/10.06	4/0.23
	23	80/17.27	24/1.09	25/1.26	4/0.11	58/12.4	22/1.01	14/0.93	1/0.04
	24	104/100.67	58/9	60/7.24	2/0.05	78/28.17	38/2.97	35/2.79	3/0.07
	25	26/5.56	8/0.48	11/0.55	0/0	18/0.64	2/0.05	0/0	0/0
	Total	836/4230.09	462/294.39	455/209.58	40/4.3	686/1074.21	345/80.6	312/62.39	23/0.94
120	18	6/0.25	0/0	1/0.02	1/0.02	4/0.92	0/0	1/0.08	0/0
	19	9/0.61	0/0	3/0.09	0/0	5/0.18	1/0.02	1/0.02	0/0
	20	14/2.82	6/0.27	7/0.39	0/0	10/2.05	6/0.18	7/0.26	1/0.02
	21	31/227.37	34/11.73	34/10.74	4/0.2	35/29.57	26/2.65	20/2.21	0/0
<i>GmFAD3C</i>	22	20/35.47	12/2.83	14/1.84	2/0.05	15/2.36	4/0.17	5/0.2	0/0
	23	12/1.13	1/0.07	8/0.34	1/0.04	3/0.09	0/0	4/0.13	0/0
	24	29/9.24	9/0.58	17/2.42	1/0.02	27/2.37	3/0.08	17/1.46	1/0.02
	25	4/0.11	1/0.02	1/0.18	1/0.02	2/0.08	1/0.03	1/0.02	0/0
	Total	125/277	63/15.5	85/16	10/0.35	101/37.62	41/3.13	56/4.39	2/0.04

(Table 2.5 continued)

	18	0/0	0/0	0/0	0/0	3/0.11	0/0	2/0.07	0/0
	19	1/0.04	4/0.09	0/0	0/0	8/0.23	2/0.06	2/0.06	0/0
	20	17/0.67	9/0.25	3/0.12	0/0	17/1.06	7/0.26	8/0.27	0/0
	21	131/9.77	31/0.88	18/0.61	0/0	154/19.32	49/2.44	27/1.39	0/0
Rice intron	22	38/1.65	122/7.08	58/3.6	1/0.02	46/3.24	155/18.27	71/7.52	0/0
	23	1/0.04	14/0.47	8/0.35	0/0	3/0.07	24/0.61	6/0.22	0/0
	24	7/0.23	16/0.46	24/1	0/0	11/0.3	41/1.25	44/2.95	0/0
	25	0/0	0/0	1/0.04	0/0	1/0.02	0/0	2/0.04	0/0
	Total	195/12.38	196/9.23	112/5.72	1/0.02	243/24.35	278/22.89	162/12.53	0/0
121	18	1/0.02	1/0.02	1/0.04	0/0	0/0	2/0.05	0/0	0/0
	19	1/0.02	0/0	0/0	0/0	2/0.12	2/0.04	0/0	0/0
	20	2/0.05	3/0.08	2/0.07	0/0	2/0.24	5/0.1	2/0.06	0/0
	21	31/1.98	17/0.44	15/0.78	0/0	43/5.81	20/0.8	11/0.43	1/0.04
Glycinin promoter	22	40/2.02	57/4.13	29/1.49	0/0	38/3.17	70/4.79	30/1.93	0/0
	23	4/0.09	7/0.18	11/0.54	0/0	6/0.28	10/0.26	6/0.18	0/0
	24	20/1.48	32/2.03	35/3.52	1/0.02	25/1.55	28/1.37	20/1.5	0/0
	25	0/0	1/0.02	3/0.07	0/0	0/0	0/0	1/0.03	0/0
	Total	99/5.66	118/6.9	96/6.5	1/0.02	116/11.16	137/7.42	70/4.12	1/0.04

(Table 2.5 continued)

	18	2/0.04	2/0.06	0/0	0/0	0/0	0/0	0/0	0/0
	19	0/0	3/0.08	0/0	0/0	2/0.04	10/0.31	2/0.04	0/0
	20	11/0.49	25/0.68	2/0.07	0/0	4/0.12	10/0.29	0/0	0/0
	21	72/7.92	63/4.07	18/0.94	0/0	61/3.35	57/2.23	17/0.74	0/0
35S	22	15/0.66	199/29.8	4/0.18	4/0.09	19/0.57	192/24.6	2/0.07	0/0
promoter	23	0/0	30/1.56	0/0	0/0	0/0	32/1.3	0/0	0/0
	24	0/0	63/3.95	6/0.22	0/0	2/0.08	56/2.73	2/0.04	0/0
	25	1/0.03	0/0	0/0	0/0	0/0	0/0	2/0.04	0/0
	Total	101/9.14	385/40.22	30/1.4	4/0.09	88/4.15	357/31.46	25/0.95	0/0
	18	350/21.48	290/12.94	193/9.66	319/18.07	0/0	0/0	0/0	0/0
	19	389/23.92	285/14.04	171/8.11	332/19.17	0/0	1/0.02	2/0.06	0/0
	20	383/24.38	311/14.61	193/9.19	341/19.19	5/0.16	6/0.13	10/0.37	5/0.15
	21	414/28.17	359/19.65	259/16.3	351/22.04	33/1.19	59/2.1	68/2.84	28/0.86
Glycinin	22	397/26.15	336/16.2	193/10.1	352/19.68	11/0.42	26/1.19	19/0.89	4/0.12
gene	23	361/21.46	283/12.65	196/8.66	334/16.61	2/0.04	1/0.03	8/0.26	0/0
	24	395/23.32	306/12.95	166/7.74	327/16.73	10/0.27	16/0.73	26/2.08	2/0.06
	25	362/18.07	303/12.75	165/6.84	336/15.56	1/0.04	2/0.04	1/0.02	0/0
	Total	3051/186.95	2473/115.79	1536/76.61	2692/147.05	62/2.12	111/4.25	134/6.52	39/1.18

(Table 2.5 continued)

	18	38/2.22	8/0.29	131/43.18	1/0.02	53/2.82	7/0.16	152/60.79	3/0.07
	19	68/4.84	6/0.16	169/116.3	0/0	84/6.02	12/0.32	211/99.13	3/0.13
	20	85/28.86	15/0.59	214/418.13	6/0.46	108/37.03	13/0.61	239/510.63	9/0.39
	21	240/405.71	52/5.95	337/4599.64	36/3.75	280/345.94	76/6.02	362/4236.21	52/3.57
Bar	22	101/40.95	91/19.76	209/1909.42	20/1.48	135/31.65	112/12.89	255/890.59	14/0.73
	23	22/2.52	16/0.65	98/87.92	1/0.02	34/1.37	19/0.59	128/64.13	1/0.04
	24	23/1.57	25/0.95	169/332.18	9/0.24	50/2.58	29/1.23	214/447.18	6/0.29
	25	3/0.11	1/0.03	57/9.58	0/0	16/0.49	2/0.04	75/11.01	0/0
	Total	580/486.78	214/28.38	1384/7516.35	73/5.97	760/427.9	270/21.88	1636/6319.65	88/5.21

123

1 Number of distinct small RNAs.

2 Total small RNA abundance.

Table 2.6 Putative functional siRNAs.

	Sequence ¹	strand	length	S-24-4D ² (CPM)	S-24-13 (CPM)	S-24-15 (CPM)	WT (CPM)
	AAUGGUGCAGAUCUUUCUGGC	-	21	1262.59	89.94	50.69	1.35
	GUAAUGGUGCAGAUCUUUCUG	-	21	88.62	6.49	4.75	0.15
	AAUGGUAAGGUGCAGAUCUU	-	21	34.45	2.67	1.36	0.04
	AUGGUGCAGAUCUUUCUGGCU	-	21	19.19	0.96	0.84	0.02
	CAGAUCUUUCUGGCUCACGGU	-	21	16.64	1.19	0.66	0
	AGAUCUUUCUGGCUCACGGUA	-	21	16.36	0.84	0.64	0
	GAUCUUUCUGGCUCACGGUAA	-	21	11.4	0.41	0.28	0.04
	GCAGAUCUUUCUGGCUCACGG	-	21	5.14	0.23	0.17	0
	AUCUUUCUGGCUCACGGUAAU	-	21	4.05	0.11	0.1	0
	UGCAGAUCUUUCUGGCUCACG	-	21	3.27	0.18	0	0
124	UGGUGCAGAUCUUUCUGGCUC	-	21	3.06	0.28	0.16	0
	UAAUGGUGCAGAUCUUUCUGG	-	21	2.79	0.14	0.13	0
	AAAUGGUAAGGUGCAGAUCU	-	21	2.25	0.24	0.11	0
	GUGCAGAUCUUUCUGGCUCAC	-	21	1.58	0.08	0.1	0
	GGUGCAGAUCUUUCUGGCUCA	-	21	1	0.04	0.04	0
	AUGGUAAGGUGCAGAUCUUU	-	21	0.97	0.04	0	0
	UCUUUCUGGCUCACGGUAAUA	-	21	0.51	0.02	0.02	0
	GAAUGGUAAGGUGCAGAUC	-	21	0.33	0	0.04	0
	UGGUAAGGUGCAGAUCUUUC	-	21	0.25	0.05	0.04	0
	GGUAAGGUGCAGAUCUUUCU	-	21	0.13	0	0	0

(Table 2.6 continued)

¹ Sequence is ranked from CPM high to low in S-24-4D.

² CPM is average of three biological replications.

LITERATURE CITED

- Alemán LM, Doench J, Sharp PA (2007) Comparison of siRNA-induced off-target RNA and protein effects. *RNA* 13 (3):385-395
- Ascherio A, Willett WC (1997) Health effects of trans fatty acids. *The American journal of clinical nutrition* 66 (4):1006S-1010S
- Axtell MJ (2013) Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol* 64:137-159. doi:10.1146/annurev-arplant-050312-120043
- Béclin C, Boutet S, Waterhouse P, Vaucheret H (2002) A branched pathway for transgene-induced RNA silencing in plants. *Curr Biol* 12 (8):684-688
- Beuselinck P, Sleper D, Bilyeu K (2006) An assessment of phenotype selection for linolenic acid using genetic markers. *Crop Sci* 46 (2):747-750
- Bilyeu K, Gillman JD, LeRoy AR (2011) Novel Mutant Allele Combinations Produce Soybeans Containing 1% Linolenic Acid in the Seed Oil. *Crop Sci* 51 (1):259. doi:10.2135/cropsci2010.01.0044
- Bilyeu K, Palavalli L, Sleper D, Beuselinck P (2003) Three microsomal omega-3 fatty-acid desaturase genes contribute to soybean linolenic acid levels. *Crop Sci* 43 (5):1833-1838
- Bilyeu K, Palavalli L, Sleper D, Beuselinck P (2005) Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. *Crop Sci* 45 (5):1830-1836
- Bilyeu K, Palavalli L, Sleper DA, Beuselinck P (2006) Molecular genetic resources for development of 1% linolenic acid soybeans. *Crop Sci* 46 (5):1913-1918

- Bleys A, Vermeersch L, Van Houdt H, Depicker A (2006) The frequency and efficiency of endogene suppression by transitive silencing signals is influenced by the length of sequence homology. *Plant Physiol* 142 (2):788-796
- Chan SW-L, Henderson IR, Jacobsen SE (2005) Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Reviews Genetics* 6 (5):351-360
- Chen PY, Weinmann L, Gaidatzis D, Pei Y, Zavolan M, Tuschl T, Meister G (2008) Strand-specific 5' -O-methylation of siRNA duplexes controls guide strand selection and targeting specificity. *RNA* 14 (2):263-274
- De Paoli E, Dorantes-Acosta A, Zhai J, Accerbi M, Jeong D-H, Park S, Meyers BC, Jorgensen RA, Green PJ (2009) Distinct extremely abundant siRNAs associated with cosuppression in petunia. *RNA* 15 (11):1965-1970
- Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O (2006) Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. *Science* 313 (5783):68-71
- Dunoyer P, Himber C, Ruiz-Ferrer V, Alioua A, Voinnet O (2007) Intra-and intercellular RNA interference in *Arabidopsis thaliana* requires components of the microRNA and heterochromatic silencing pathways. *Nat Genet* 39 (7):848-856
- Eamens A, Wang MB, Smith NA, Waterhouse PM (2008) RNA silencing in plants: yesterday, today, and tomorrow. *Plant Physiol* 147 (2):456-468. doi:10.1104/pp.108.117275
- Flores T, Karpova O, Su X, Zeng P, Bilyeu K, Sleper DA, Nguyen HT, Zhang ZJ (2008) Silencing of GmFAD3 gene by siRNA leads to low α -linolenic acids (18: 3) of fad3-mutant phenotype in soybean [*Glycine max* (Merr.)]. *Transgenic Res* 17 (5):839-850
- Frizzi A, Huang S (2010) Tapping RNA silencing pathways for plant biotechnology. *Plant Biotechnol J* 8 (6):655-677
- Gascioli V, Mallory AC, Bartel DP, Vaucheret H (2005) Partially Redundant Functions of *Arabidopsis* DICER-like Enzymes and a Role for DCL4 in Producing *trans*-Acting siRNAs. *Curr Biol* 15 (16):1494-1500

- Hamilton A, Voinnet O, Chappell L, Baulcombe D (2002) Two classes of short interfering RNA in RNA silencing. *The EMBO Journal* 21 (17):4671-4679
- Hu FB, Stampfer MJ, Manson JE, Rimm E, Colditz GA, Rosner BA, Hennekens CH, Willett WC (1997) Dietary fat intake and the risk of coronary heart disease in women. *New Engl J Med* 337 (21):1491-1499
- Huettel B, Kanno T, Daxinger L, Bucher E, van der Winden J, Matzke AJ, Matzke M (2007) RNA-directed DNA methylation mediated by DRD1 and Pol IVb: a versatile pathway for transcriptional gene silencing in plants. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression* 1769 (5):358-374
- Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol* 21 (6):635-637
- Jackson AL, Burchard J, Leake D, Reynolds A, Schelter J, Guo J, Johnson JM, Lim L, Karpilow J, Nichols K (2006) Position-specific chemical modification of siRNAs reduces “off-target” transcript silencing. *RNA* 12 (7):1197-1205
- Khvorova A, Reynolds A, Jayasena SD (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115 (2):209-216
- Kościańska E, Kalantidis K, Wypijewski K, Sadowski J, Tabler M (2005) Analysis of RNA silencing in agroinfiltrated leaves of *Nicotiana benthamiana* and *Nicotiana tabacum*. *Plant Mol Biol* 59 (4):647-661
- Kusaba M (2004) RNA interference in crop plants. *Curr Opin Biotechnol* 15 (2):139-143
- Kusaba M, Miyahara K, Iida S, Fukuoka H, Takano T, Sassa H, Nishimura M, Nishio T (2003) Low glutelin content1: a dominant mutation that suppresses the glutelin multigene family via RNA silencing in rice. *The Plant Cell Online* 15 (6):1455-1467
- Liu H-R, White PJ (1992) Oxidative stability of soybean oils with altered fatty acid compositions. *Journal of the American Oil Chemists Society* 69 (6):528-532

- Matranga C, Tomari Y, Shin C, Bartel DP, Zamore PD (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* 123 (4):607-620
- Miki D, Itoh R, Shimamoto K (2005) RNA silencing of single and multiple members in a gene family of rice. *Plant Physiol* 138 (4):1903-1913
- Mishiba Ki, Nishihara M, Nakatsuka T, Abe Y, Hirano H, Yokoi T, Kikuchi A, Yamamura S (2005) Consistent transcriptional silencing of 35S - driven transgenes in gentian. *The Plant Journal* 44 (4):541-556
- Mlotshwa S, Pruss GJ, Gao Z, Mgutshini NL, Li J, Chen X, Bowman LH, Vance V (2010) Transcriptional silencing induced by Arabidopsis T - DNA mutants is associated with 35S promoter siRNAs and requires genes involved in siRNA - mediated chromatin silencing. *The Plant Journal* 64 (4):699-704
- Mlotshwa S, Pruss GJ, Peragine A, Endres MW, Li J, Chen X, Poethig RS, Bowman LH, Vance V (2008) DICER-LIKE2 plays a primary role in transitive silencing of transgenes in Arabidopsis. *PLoS One* 3 (3):e1755
- Mourrain P, van Blokland R, Kooter JM, Vaucheret H (2007) A single transgene locus triggers both transcriptional and post-transcriptional silencing through double-stranded RNA production. *Planta* 225 (2):365-379
- Ogita S, Uefuji H, Yamaguchi Y, Koizumi N, Sano H (2003) RNA interference: producing decaffeinated coffee plants. *Nature* 423 (6942):823-823
- Petersen BO, Albrechtsen M (2005) Evidence implying only unprimed RdRP activity during transitive gene silencing in plants. *Plant Mol Biol* 58 (4):575-583
- Qiu S, Adema CM, Lane T (2005) A computational study of off-target effects of RNA interference. *Nucleic Acids Res* 33 (6):1834-1847
- Rand TA, Petersen S, Du F, Wang X (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* 123 (4):621-629

- Sanders M, Maddelein W, Depicker A, Van Montagu M, Cornelissen M, Jacobs J (2002) An active role for endogenous β - 1, 3 - glucanase genes in transgene - mediated co - suppression in tobacco. *The EMBO journal* 21 (21):5824-5832
- Schwarz DS, Hutvágner G, Du T, Xu Z, Aronin N, Zamore PD (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115 (2):199-208
- Senthil - Kumar M, Hema R, Anand A, Kang L, Udayakumar M, Mysore KS (2007) A systematic study to determine the extent of gene silencing in *Nicotiana benthamiana* and other Solanaceae species when heterologous gene sequences are used for virus - induced gene silencing. *New Phytol* 176 (4):782-791
- Shoemaker R, Polzin K, Labate J, Specht J, Brummer E, Olson T, Young N, Concibido V, Wilcox J, Tamulonis J (1996) Genome duplication in soybean (*Glycine subgenus soja*). *Genetics* 144 (1):329-338
- Small I (2007) RNAi for revealing and engineering plant gene functions. *Curr Opin Biotechnol* 18 (2):148-153
- Stoutjesdijk PA, Singh SP, Liu Q, Hurlstone CJ, Waterhouse PA, Green AG (2002) hpRNA-mediated targeting of the *Arabidopsis* FAD2 gene gives highly efficient and stable silencing. *Plant Physiol* 129 (4):1723-1731
- Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y (2008) The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. *Plant and cell physiology* 49 (4):493-500
- Tang G, Reinhart BJ, Bartel DP, Zamore PD (2003) A biochemical framework for RNA silencing in plants. *Genes Dev* 17 (1):49-63
- Tosar JP, Rovira C, Naya H, Cayota A (2014) Mining of public sequencing databases supports a non-dietary origin for putative foreign miRNAs: underestimated effects of contamination in NGS. *RNA*
- Vaistij FE, Jones L, Baulcombe DC (2002) Spreading of RNA targeting and DNA methylation in RNA silencing requires transcription of the target gene and a putative RNA-dependent RNA polymerase. *The Plant Cell Online* 14 (4):857-867

Watson JM, Fusaro AF, Wang M, Waterhouse PM (2005) RNA silencing platforms in plants. *FEBS Lett* 579 (26):5982-5987

Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, Jacobsen SE, Carrington JC (2004) Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol* 2 (5):e104

Yadav NS, Wierzbicki A, Aegerter M, Caster CS, Perez-Grau L, Kinney AJ, Hitz WD, Booth Jr JR, Schweiger B, Stecca KL (1993) Cloning of Higher Plant [ω]-3 Fatty Acid Desaturases. *Plant Physiol* 103 (2):467-476

CHAPTER 3

Development of atasiRNA-based Vectors for Efficient Gene Silencing in Soybean

ABSTRACT

In this study, we developed two *Arabidopsis TAS1a*-based atasiRNA constructs targeting the *GmFAD3* gene family using online siRNA design tool OligoWalk. However, computational predicted siRNAs does not represent their *in vivo* efficacy. Further investigation is needed to determine whether siRNA candidates could conduct efficient silencing of target genes in plants. The *Agrobacterium*-mediated transformation results may reveal an adverse effect of miR173 overexpression during soybean transformation as well as the instability of RNAi. For the two resultant transgenic RNAi lines, elevated linoleic acid content was obtained, but the proof that such an increase was achieved through the silencing of *FAD3* genes still need further exploration. A transient expression assay through leaf-infiltration of both cDNA and atasiRNA binary constructs carried within *Agrobacterium* strains was developed to validate the silencing efficacy of candidate atasiRNAs, but the down-regulation of the target gene was not as efficient as in stably transformed soybean plants. Thus, we suggest to incorporate both cDNA and atasiRNA expression cassette into a single binary vector, to minimize variations caused by separate constructs. Furthermore, to simplify the deployment of atasiRNA platform and investigate the utility of miR390 and TAS3 as a gene silencing tool in soybean, spacial and temporal analysis of miR390 was performed. Our results implicated that miR390 is consistently expressed in all sampled tissues with the highest abundance in flowers and early stage of pod development, which makes miR390 a good candidate to trigger the formation of atasiRNA in soybean.

INTRODUCTION

Endogenous non-coding small RNAs are important regulators of gene expression that act at the transcriptional or post-transcriptional level. In recent years, a novel class of 21-22nt regulatory siRNA termed *trans*-acting small interfering RNA (tasiRNA) has been identified and intensively studied for their biogenesis and functions (Peragine et al. 2004; Hunter et al. 2006; Montgomery et al. 2008; Williams et al. 2005; Xie et al. 2005b; Yoshikawa et al. 2005; Garcia et al. 2006; Adenot et al. 2006; Allen et al. 2005; Vazquez et al. 2004)

The generation of tasiRNAs is triggered by miRNA-guided cleavage of tasiRNA-generating (*TAS*) gene derived transcript. One of the 3' or 5' cleavage products is converted to dsRNA by RDR6 and subsequently processed by DCL4 into 21-nt siRNAs that are phased with regard to the miRNA cleavage site (Peragine et al. 2004; Allen et al. 2005; Xie et al. 2005b; Yoshikawa et al. 2005). One strand of the tasiRNA duplex is selectively assembled into RNA-induced silencing complex (RISC) to catalyze the cleavage or repress the translation of complementary target mRNAs. The identification of tasiRNAs establishes a link between the miRNA and siRNA pathways and also raises the possibility that numerous endogenous genes may be regulated by this special subgroup of siRNAs.

Arabidopsis thaliana contains eight *trans*-acting siRNA (*TAS*) loci belonging to four families. *TAS1* and *TAS2* were the first characterized *TAS* transcripts targeted by miR173, and these non-protein-coding genes are only found in *Arabidopsis* (Allen et al.

2005; Vazquez et al. 2004; Yoshikawa et al. 2005; Vaucheret 2005). tasiRNAs from *TAS4* locus was identified much later than *TAS1/2* due to low tasiRNA abundance (Rajagopalan et al. 2006). Unlike *TAS1/2*, the *TAS4* transcript is a target of miR828 and is known to exist in dicots (Rajagopalan et al. 2006). *TAS3* is conserved among higher plant species and its tasiRNAs originate from sequences between two miR390 binding sites (Montgomery et al. 2008). There are three *TAS3* loci in *Arabidopsis*: *TAS3a* (At3g17185), *TAS3b* (At5g49615), and *TAS3c* (At5g57735) (Howell et al. 2007). *TAS1*, *TAS2*, *TAS4* only require one miRNA binding site upstream of the tasiRNA-generating region; while *TAS3* require two miR390 binding sites, one upstream and one downstream of the tasiRNA-generating region. These two different classes each requires unique components and unique RNA structural features for their biogenesis of tasiRNAs.

In recent years, artificial tasiRNA (ataiRNA)-directed gene silencing has been developed and gained significant attention in plants (Felippes and Weigel 2009; de Felippes et al. 2011; Zhang 2014; Feng et al. 2013; de la Luz Gutiérrez-Nava et al. 2008; Ossowski et al. 2008; Carbonell et al. 2014). In an ataiRNA inducing vector, the transgene is simply made from a modified *TAS* gene sequence by substituting a single or several copies of native siRNA with ataiRNAs targeting particular genes (Ossowski et al. 2008; Felippes and Weigel 2009; de Felippes et al. 2011; de la Luz Gutiérrez-Nava et al. 2008; Carbonell et al. 2014). Compared with hpRNA, the ataiRNA approach is more practical for stacking multiple functional small RNA sequences into a single construct and minimizing the possibility of off-target effects (Ossowski et al. 2008; Zhang 2014). However, most of the

studies on atasiRNA-mediated gene silencing has been done in model plants with relatively simple genomes.

Soybean (*Glycine max*) ranks among the top oil and protein crops in the U.S.. One of the most important goals of oil quality breeding in soybean has been to lower its α -linolenic acid (18:3) content for improved oxidative stability and flavor to eliminate the need for hydrogenation. In soybean developing seed, α -linolenic acid (18:3) is produced from linoleic acid (18:2) under the catalytic activity of omega-3 fatty acid desaturase (*FAD3*) enzyme. Inhibition of *FAD3* in soybeans reduces the level of unstable linolenic acid (18:3) and the resultant soybean oil can be directly used without hydrogenation. Three independent microsomal omega-3 acid desaturases (*GmFAD3A*, *GmFAD3B*, *GmFAD3C*) have been characterized and linked to low seed linolenic acid phenotype in soybean (Bilyeu et al. 2003). Of the three, *GmFAD3A* (Glyma14g37350) has been shown to be the major contributors to seed linolenic acid levels (Bilyeu et al. 2003; Bilyeu et al. 2005; Bilyeu et al. 2006; Bilyeu et al. 2011). *GmFAD3B* (Glyma02g39230) and *GmFAD3C* (Glyma18g06950) shares 94% and 79% sequence similarity with *GmFAD3A* in the coding regions, respectively. These two genes have a less impact on seed linolenic acid levels (Bilyeu et al. 2003; Bilyeu et al. 2005; Bilyeu et al. 2006; Bilyeu et al. 2011).

This research has been undertaken to employ atasiRNA technology as a tool to analyze gene function and manipulate commercial traits in soybean. To fully capitalize on the potential of atasiRNA-mediated silencing, the *GmFAD3* gene family was chosen as a test model for our gene silencing study. The primary goal of this work was to validate and further improve atasiRNA-mediated RNAi as a gene silencing strategy in soybean for

potent and specific gene silencing with desired trait modification(s). Our long-term goal is to achieve gene stacking through atasiRNA.

MATERIALS AND METHODS

Plant material and growth conditions

Soybean elite genotype “Maverick” (2010) was obtained from the Missouri Foundation Seeds (Portageville, MO) and used for development of transgenic soybean lines. All soybean were grown on Pro-mix soil (SunGro, Agawam, MA) in 13-litter pots in a greenhouse under controlled-environmental conditions at 23-26°C with supplemental 50-90 Klux day light intensity and 12/12 h photoperiod from late May to early November or a 14/10 h photoperiod during the rest seasons. Plants were fertilized once with Osmocote 14-14-14 (Hummert International, Earth City, MO) at the time of planting or transplanting and watered as needed.

RNAi vector construction

The base vector pUB14 for all ata-siRNA constructs was assembled. One key feature of this vector is that harbored two separate T-DNAs. T-DNA one contained multiple cloning sites for the insertion of the gene of interest, while T-DNA two carried the *bar* (bialaphosresistance) gene expression cassette as a plant selection marker. OligoWalk (http://rna.urmc.rochester.edu/cgi-bin/server_exe/oligowalk/oliGowalkform.cgi) was used for the design and quality control of the 21-bp artificial tasiRNA (Table 3.1). *Arabidopsis*

TAS1a transcript (At2g27400) was modified by substituting the native siR255 sequence with an atasiRNA sequence as described by Felippes et al. (Felippes and Weigel 2009). The 2x 35S or soybean Glycinin promoter (Flores et al. 2008a) and the soybean vegetative storage protein gene terminator (Tvsp) (Rhee and Staswick 1992) were added at the 5' and 3' end of the modified *TAS1a* sequence as well as a sequence containing a foldback structure of *Arabidopsis MIR173* (Chen et al. 2010), respectively. The resultant two cassettes were synthesized (Genscript, Piscataway, NJ) and cloned sequentially as *EcoRI-BamHI/BamHI-PstI* fragments into the T-DNA one region of pUB14 and the resultant vector was introduced into *A. tumefaciens* strain AGL1 by the freeze-thaw method (Figure 3.1) (Chen et al. 1994).

cDNA overexpression vector construction

Full length cDNAs of *GmFAD3A* (Glyma14g37350), *GmFAD3B* (Glyma02g39230) and *GmFAD3C* (Glyma18g06950) were cloned from soybean genotype “Jack” and ligated to pCR®8/GW/TOPO® TA vector using pCR®8/GW/TOPO® TA Cloning Kit (Invitrogen). The cDNA sequence, now in the resultant entry clone, was subsequently introduced into pEarleyGate203 vector via Gateway recombination cloning using Gateway® LR Clonase® enzyme mix (Invitrogen) (Earley et al. 2006).

Plant transformation

Soybean transformation followed the *Agrobacterium*-mediated cotyledonary-node transformation system as previously described (Zhang et al. 1999a; Zeng et al. 2004; Wright et al. 2010). Major modifications include the addition of “dip-wounding” during the explant preparation. In this procedure, explants were prewounded before cocultivation with blade dipped in *Agrobacterium* suspension to increase transformation frequency (Barampuzam and Zhang 2011). In addition, a revised selection scheme of glufosinate was applied at 0, 10, and 4mg/L during first and second shoot initiation and shoot elongation stages, respectively.

Transgene integration and segregation analysis

Regenerated plants were leaf-painted three times with 200mg/L Liberty Herbicide (Bayer CropScience, Monheim am Rhein, Germany) to screen for putative transformants as previously described (Zhang et al. 1999b). All of the screened plants were then sampled and analyzed by PCR to verify the insertion of the *bar* gene and/or the gene of interest using the REDExtract-N-Amp™ Plant PCR kit (Sigma-Aldrich, St. Louis, MO). The parameters for the PCR were as follows, 94°C for 3min, then 35 cycles of 30s denaturation at 94°C, 30s annealing at 60°C, 30s extension at 72°C, and 10min final extension at 72°C. Primer information is listed in Table 3.2. Putative transgenic soybean lines were subsequently subjected to progeny segregation analysis by leaf-painting assay and PCR analysis to determine the inheritance of transgenes from T₀ plants.

Fatty acid analysis

The fatty acid profiles of dry mature soybean seeds from transgenic and wild type control samples were examined by a gas chromatography (GC) method as previously described (Beuselinck et al. 2006). Ten seeds from each plant were crushed in an envelope and used for fatty acid determination. For each transgenic soybean event, seeds from three plants were individually analyzed. The individual fatty acid contents of palmitic, stearic, oleic, linoleic, and linolenic acids are represented as a proportion of total fatty acids in the extracted oil.

Agroinfiltration

Wild type *Nicotiana benthamiana* plants were grown from seed and maintained in a growth chamber under a 16 h light/8 h dark photoperiod at 25°C. Four week old plants were used for the infiltration assay.

For agroinfiltration, all T-DNA constructs were introduced into *A. tumefaciens* GV3101 using the freeze-thaw method (Chen et al. 1994). Recombinant *A. tumefaciens* was grown overnight at 28 °C in 10ml LB liquid medium containing 50 mg/L kanamycin, 10mg/L rifampicin and 20 µM acetosyringone to an OD₆₀₀=1.0. Cells were harvested by centrifugation and resuspended to a final concentration of OD₆₀₀=0.6 in 10mM MgCl₂, 10mM MES, pH 5.6, 150 µM acetosyringone. The cell suspension was incubated

for 5h at room temperature before infiltration. *A. tumefaciens* cultures containing the atasiRNA or control vector and target gene-myc construct were mixed at a 1:1 ratio. The mixtures were co-infiltrated into leaves of *N. bethamiana* plants using a 1ml syringe with no needle, and the infiltrated area was outlined with a marker pen. For each construct, a total of 6 leaves on two *N. bethamiana* plants were infiltrated.

qRT-PCR

The marked area on each infiltrated *N. bethamiana* leaf was collected 60 hrs after infiltration. For each construct, total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA) from a bulk sample of 6 infiltrated leaves from two individual plants.

One-step real-time quantitative PCR was performed in triplicate biological and technical replications on an CFX-96™ Real-Time system (Bio-Rad, Hercules, CA) with the recommended settings for SYBR Green. Each reaction contained 2 µl DNase I (Invitrogen, Carlsbad, CA) treated RNA, 10µM of each specific primer, 0.25 µl iScript Reverse Transcriptase and 10µl of 2x SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad) in a final volume of 20 µl. Genomic DNA and other contamination were monitored by no-template and no-RT controls. A standard curve was generated from cDNAs to determine the PCR efficiency of each primer pair. The following PCR program was used for all PCR reactions: 50 °C for 10 min, 95°C for 1 min, followed by 40 cycles

of 10s denaturation at 95°C, 30s annealing and extension at 60°C. Amplification specificity was verified by melting curve analysis at the end of PCR.

Templates were normalized for difference in RNA amount using L23 amplification levels (Liu et al. 2012). Data were analyzed with BioRad CFX Manager™ 2.0 Software (Bio-Rad). The comparative threshold cycle method ($\Delta\Delta C_t$) was used to determine relative transcript abundance levels. Sequences of applied primers are listed in Table 3.3.

stem-loop RT-PCR

The expression profile of mature miR390a was assayed by stem-loop RT-PCR and performed as previously described (Varkonyi-Gasic et al. 2007). The Gma-miR390a mature sequence (Accession NO. MI0007214) was obtained from miRBase database (<http://www.mirbase.org>). Stem-loop RT primers were designed according to previous protocols (Varkonyi-Gasic et al. 2007; Chen et al. 2005) with a six-nucleotide extension complementary to the 3' end of mature miRNA and a Universal ProbeLibrary Probe #21 sequence binding site. cDNA was synthesized from 500ng DNase I (Invitrogen) treated total RNA using Superscript III reverse transcriptase (Invitrogen) with miRNA specific stem-loop primers. The reaction was incubated at 16°C for 30 min, followed by pulsed RT of 60 cycles at 30°C for 30s, 42°C for 30s and 50°C for 1s and finally 5 min at 85°C. 2µl of cDNA product from each sample were subjected to quantitative real-time PCR analysis using a miRNA-specific forward primer and a universal reverse primer. Each 20µl PCR

reaction contained 10 μ M forward primer, 10 μ M reverse primer, 5 μ M UPL probe #21 and 1 \times Bio-rad SsoFast Probes Supermix (Bio-Rad). The PCR reaction was performed in triplicate biological and technical replications on an CFX-96TM Real-Time system (Bio-Rad) at 95 $^{\circ}$ C for 5min, followed by 45 cycles of 95 $^{\circ}$ C for 5s, 60 $^{\circ}$ C for 10s, and 72 $^{\circ}$ C for 1s.

PCR efficiency of each primer pair was determined by standard curves based on a fivefold dilution series. Data were analyzed with BioRad CFX ManagerTM 2.0 Software (Bio-Rad). The comparative threshold cycle method ($\Delta\Delta$ Ct) was used to determine relative transcript abundance levels. 5.8S rRNA was used as the inner control. Sequences of stem–loop RT primers, miRNA-specific PCR primers and universal primers are listed in Table 3.4.

RESULTS

Comparison of atasiRNA candidates designed by OligoWalk with putative functional siRNAs experimentally determined by 5' RACE

In our previous study, a hpRNA-based RNAi vector pMUFAD was designed to effectively silence the three active members of soybean *FAD3* gene family (Flores et al. 2008b). To ensure efficient target silencing and compare atasiRNA strategy with hpRNA-mediated RNAi, the 318-nt IR of pMUFAD was used to generate atasiRNA candidates using the online siRNA design tool OligoWalk (http://rna.urmc.rochester.edu/cgi-bin/server_exe/oligowalk/oligowalk_form.cgi) (Lu and Mathews 2008). A total of 78 siRNA candidates were generated for the 318-bp target and computationally ranked by sequence complementarity and hybridization energy (Table 3.1). At that time, we chose a 21-nt siRNA candidate starting at position 149 with a 90.54% possibility of being efficient siRNA from the top of the OligoWalk output list to construct the atasiRNA construct. This siRNA candidate shares 100% sequence homology with *GmFAD3A* and *GmFAD3B*, and exhibited one mismatch with *GmFAD3C* at position 4 from the 5' end of siRNA sequence.

As shown in Table 3.1, more than 50% of the siRNA candidates designed by OligoWalk are located at the first half of the 318-nt IR. Especially, 9 out of 10 top candidates showed a starting position prior to nucleotide 159. When compared with the 5' RACE result conducted in the previous chapter, only two out of the 78 listed siRNA

candidates cover one or both of the two predominant cleavage sites (297-298, 301-302). The possibility for these two candidates being an efficient siRNA is 71.99% and 69.97%, ranked 42th and 49th on the putative siRNA list, respectively. The inconsistency between OligoWalk designed siRNA candidates and experimentally determined cleavage sites may reflect that the OligoWalk ranking could not accurately predict optimal siRNAs *in vivo*. However, the cleavage events directed by 318-bp siRNAs were also affected by their relative abundance when processed from the hpRNA precursor. Thus, whether siRNA candidates designed by OligoWalk could deliver efficient silencing of target genes still need further investigation.

***Agrobacterium*-mediated transformation did not yield enough transgenic plants**

For *Agrobacterium*-mediated transformation, a total of 3,627 explants were inoculated with 4 different plasmids (Table 3.5). The two control vectors, 35S-173 and GlyP-173, contain a miR173 expression cassette under the control of 35S and Glycinin promoter, respectively, while the two atasiRNA vectors, 35S-FAD and GlyP-FAD, contain both miR173 and modified *TAS1a* expression cassette with one atasiRNA targeting *GmFAD3*.

The percentage of good regeneration ranged from 20.6% to 30.24% and shoot regeneration frequency was around 4%, which seemed normal. However, only less than half of the shoots survived till greenhouse stage. Of the four plasmids, 35S-FAD exhibited

highest number of putative transgenic plants based on leaf painting results. Surprisingly, all of the 21 regenerated plants from GlyP-173 showed a susceptible phenotype after leaf painting. Because a two T-DNA vector was used to as a backbone for these constructs, all regenerated plants were PCR screened for the presence of *bar* and *TAS1a* cassette. Instead of an independent insertion mode, PCR results implicated that insertion of the two T-DNAs were tightly linked. For the 15 transgenic T₀ plants, only 3 of them passed the transgene to the next generation. No control plants were obtained at this point.

Fatty acid profile revealed increased linoleic acid content in two of the analyzed RNAi lines

The full fatty acid profiles of the seeds of the three RNAi lines were analyzed using gas chromatography. 10 T₂ seeds from three individual plants of three hemizygous *FAD3* RNAi lines were sampled. Individual fatty acid contents are presented as the relative percent of palmitic, stearic, oleic, linoleic, and linolenic acids from the total amount of extracted oil. Although only one of the nine transgenic plants displayed a significant reduction in linolenic acid content compared to that of WT control “Maverick” ($\alpha=0.05$), suggesting a potent silencing of *GmFAD3* (Table 3.6), all six plants containing 35S-FAD accumulated significantly higher linoleic acid levels in the seed oil than WT control, ranging from 59.4% to 60.5% ($\alpha=0.05$, Table 3.6). This effect might result from a decrease in seed expressed *FAD3* enzyme activity responsible for converting linoleic acid precursors into linolenic acid precursors. By contrast, all of the three transgenic plants containing

GlyP-FAD displayed similar linoleic acid and higher linolenic content of WT, indicating that the atasiRNA expression cassette driven by glycinin promoter did not work efficiently or was not inherited. However, due to the lack of control plants only expressing miR173, further investigation is needed to determine whether elevated linoleic acid content observed in this study was achieved through the silencing of *FAD3* genes. It is possible such increase was induced by the expression of miR173, which does not exist in soybean.

Test silencing efficacy of atasiRNA constructs using transient expression assay

Since stable transformation experiment did not yield enough transgenic plants, a transient expression assay using agroinfiltrated *N. benthamiana* plants were developed to test the silencing efficacy of atasiRNA constructs designed by OligoWalk. Prior to infiltration a suspension of *Agrobacterium* harboring cDNA plasmids of *GmFAD3A*, *GmFAD3B*, and *GmFAD3C* was mixed in a 1:1 ratio with a GV3101 suspension as positive controls. Suspensions of *Agrobacterium* harboring atasiRNA plasmids, pMUFAD, and control plasmids were mixed in a 1:1 ratio with a bacterial suspension carrying corresponding cDNA plasmids. GV3101 infiltrated and WT *N. benthamiana* leaf samples were used as negative control. To reduce experiment variation, a bulk of 6 infiltrated leaves from 2 individual plants were sampled for each treatment. Relative target gene cDNA levels in infiltrated *Nicotiana benthamiana* leaves were detected using qRT-PCR.

As shown in Figure 3.2, the cDNA positive control always had the highest expression among all treatments, suggesting that mix of *Agrobacterium* carrying two different plasmids may affect the expression of each other in infiltrated *Nicotiana benthamiana* leaves. All of the RNAi vectors exhibited lower target transcript levels when compared to the corresponding control vector, indicating the potential silencing of target genes. However, the differences between control and atasiRNA/hpRNA treatment were not significant except pMUFAD targeting *FAD3C*. This result is inconsistent with that obtained in stably transformed soybean plants of pMUFAD, where the silencing of *FAD3A* and *FAD3B* was much efficient than *FAD3C*. Since not all plasmids share the same expression levels in infiltrated *Nicotiana benthamiana* leaves, the inefficient silencing observed in this experiment might be due to the low expression level of RNAi vector relative to that of cDNA vector. Besides, the transcript level of over expressed cDNAs are much higher than that of endogenous *FAD3* genes, which might make efficient target silencing more difficult.

MiR390 temporal and spacial expression analysis

Our soybean stable transformation experiment using atasiRNA constructs did not yield good results. One possibility is that *Arabidopsis* miR173 might have an adverse effect on regeneration. As mentioned previously, tasiRNAs generated from TAS3 locus is triggered by miR390, which eliminate the need of miR173 as an inducer. Thus, atasiRNA approach employing miR390 and TAS3 seems a good alternative as a gene silencing tool in soybean.

miR390 is conserved among plant species, which indirectly regulate a family of Auxin Response Factor that impact the development, growth, and physiology of plants (Garcia 2008; Marin et al. 2010). According to the literature, overexpression, of essential miRNAs could lead to developmental abnormalities (Garcia 2008). Introduction of a strongly expressed extra copy of miR390 into soybean might be hazardous. Thus, it is important to obtain a comprehensive understanding of endogenous soybean miR390 expression pattern in different tissue and development stage. Two miR390 loci was found in *Arabidopsis*, *ath-MIR390a* (miRBase: MI0001000) and *ath-MIR390b* (miRBase: MI0001001), both of which share the same mature miRNA sequence (Xie et al. 2005a). Soybean possesses 7 *MIR390* loci, but only *gma-MIR390a* (miRBase: MI0007214), *gma-MIR390f* (miRBase: MI0021702) and *gma-MIR390g* (miRBase: MI0021703) produce mature miR390 identical to that in *Arabidopsis* (Subramanian et al. 2008; Joshi et al. 2010; Radwan et al. 2011; Turner et al. 2012).

In this study, the abundance of mature miR390 identical to the one in *Arabidopsis* was determined using a stem-loop RT-PCR approach (Varkonyi-Gasic et al. 2007; Chen et al. 2005). A total of 9 different development stages were selected for the temporal expression analysis and 8 different tissue types were chosen for the spacial expression analysis. The abundance of miR390 in each sample was determined relative to that in one day germinating seed (Figure 3.3). miR390 is consistently expressed in all sampled tissues. Leaf samples exhibited overall higher abundance than that in stems and roots. The expression of miR390 in vegetative tissue exhibited no significant changes over time, with slightly higher abundance in seedlings. In contrast, high abundant miR390 in flowers and

young pods were detected, implicating that miR390 are actively involved in regulatory pathways specifically within these two types of tissues, while the expression of miR390 remains relatively low in developing seeds.

DISCUSSION

In this study, we developed two *Arabidopsis TAS1a*-based atasiRNA constructs targeting the *GmFAD3* gene family. The atasiRNA candidates were generated from the 318-bp IR of pMUFAD using online siRNA design tool OligoWalk (http://rna.urmc.rochester.edu/cgi-bin/server_exe/oligowalk/oligowalk_form.cgi) (Lu and Mathews 2008). However, experimentally determined siRNA-mediated cleavage sites in the previous chapter suggested different functional siRNAs from OligoWalk predicted optimal siRNAs. It is possible that the OligoWalk ranking using sequence complementarity and hybridization energy does not represent siRNA's *in vivo* efficacy. In a recent study, Li et al. tested the amiRNA efficacy designed by another web-based tool WMD (<http://wmd3.weigelworld.org/cgi-bin/webapp.cgi>), and found the same problem with computational predicted amiRNAs (Li et al. 2013). Thus, further investigation is needed to determine whether siRNA candidates designed by OligoWalk could conduct efficient silencing of target genes in plants. Meanwhile, further design of atasiRNA may not be limited to the 318-bp IR region to further secure effective silencing of *GmFAD3*.

Agrobacterium-mediated transformation using atasiRNA constructs did not yield enough transgenic plants. Although all experiments produced the normal amount of regenerated plants, most of them are escapes with no T-DNA insertions. Such a high escape rate is intriguing, which may suggest an adverse effect of miR173 during soybean transformation. For those PCR-confirmed transgenic T₀ plants, only 20% of them passed

the transgene to the next generation. This result raised our concern about the instability of miR173-mediated RNAi, which might need further exploration before applying this technology for the genetic improvement of crop plants.

The full fatty acid profiles of two RNAi lines containing 35S-FAD implicated significantly increased linoleic acid levels in the seed oil than WT control. This effect might result from the silencing of *GmFAD3* which blocked the conversion of linoleic acid precursors to linolenic precursors in the soybean seed. While all of the three transgenic plants containing GlyP-FAD displayed similar linoleic acid and higher linolenic content of WT, indicating that the atasiRNA expression cassette driven by glycinin promoter did not work efficiently or was not inherited. However, due to the lack of control plants only expressing miR173, further investigation is needed to determine whether elevated linoleic acid content observed in this study was achieved through the silencing of *FAD3* genes. In any case we noticed that silencing potency of 35S-FAD transgenic plants (this study) were much less than that of pMUFAD transgenic counterparts (Chapter II), suggesting that either most effective atasiRNA candidate remains to be selected and deployed or the atasiRNA platform is intrinsically less potent than hpRNA.

Due to the unsuccessful stable transformation experiment, we developed a transient expression assay to test the silencing efficacy of atasiRNA constructs. By using a bulk of 6 infiltrated leaf samples from 2 individual plants, technical variations were successfully restricted within a small range. However, the down-regulation of target gene was not as efficient as in stably transformed soybean plants. Since different plasmids may exhibit variant expression levels in infiltrated *Nicotiana benthamiana* leaves, the inefficient

silencing might be due to the low expression level of RNAi vector relative to that of cDNA vector. Besides, the transcript level of over expressed cDNAs are much higher than that of the endogenous *FAD3* genes, which might make efficient target silencing more difficult. Thus, for the comparison of atasiRNA efficacy, it would be better to incorporate both cDNA and atasiRNA expression cassette into the same binary vector, to eliminate any possibility variations caused by expression differences.

In order to investigate whether miR390 and TAS3 could be employed as a gene silencing tool in soybean, spatiotemporal expression of soybean miR390 was investigated. Our results implicated that miR390 is consistently expressed in all sampled tissues with the highest abundance in flowers and young pods, which makes miR390 a good candidate to trigger the formation of atasiRNA in soybean.

FIGURES

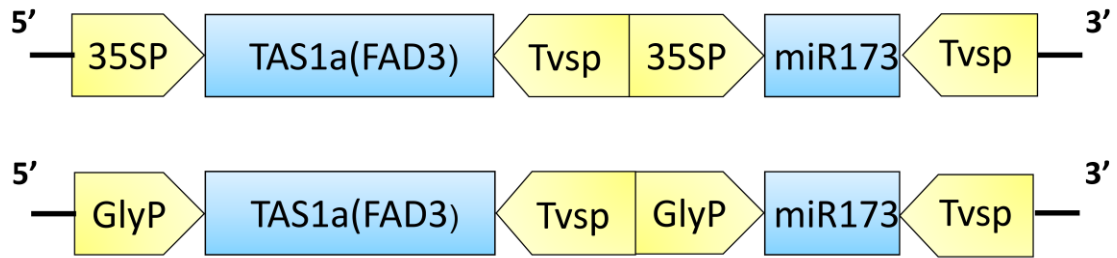


Figure 3.1 Schematic presentation of the atasiRNA-producing region of 35S-FAD and GlyP-FAD. Tvsp represent soybean vegetative storage protein gene terminator; 35SP stand for Cavliflower mosaic virus 35S promoter; GlyP stand for soybean glycinin gene promoter; miR173 indicate precursor sequence of Arabidopsis *MIR173*; TAS1a represent the Arabidopsis *TAS1a* transcript (At2g27400); *FAD3*, 21-nt *GmFAD3* target sequence substitute the native siR255 sequence.

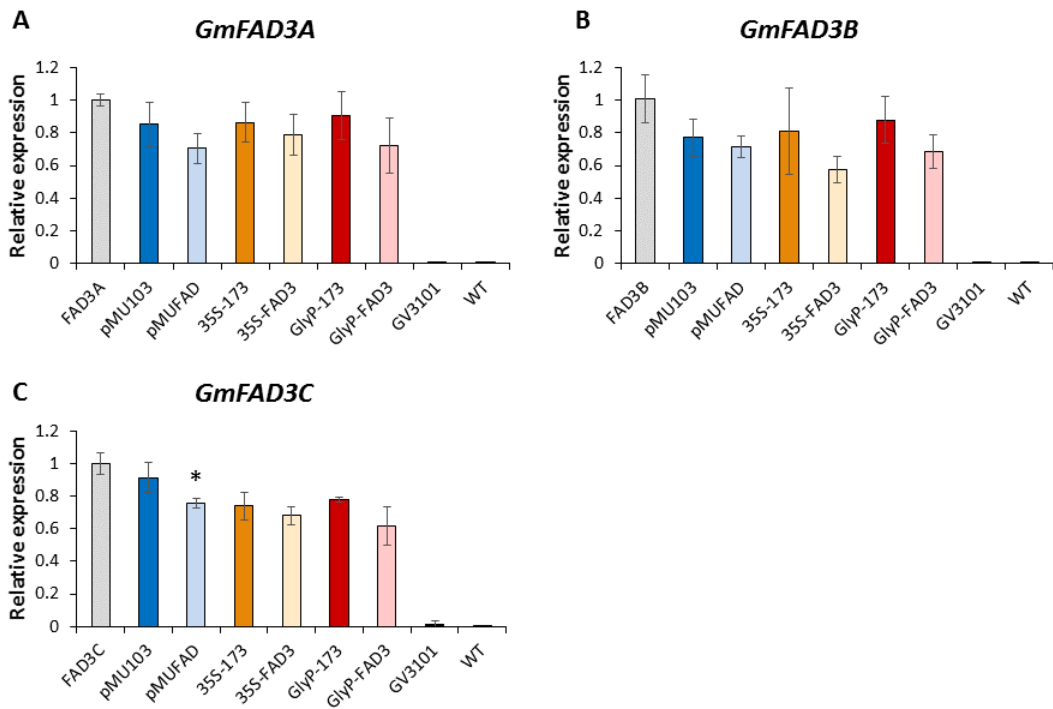


Figure 3.2 Relative target gene cDNA levels in agro-infiltrated *Nicotiana benthamiana* plants. Data are averages of biological triplicates \pm SD normalized to L23 mRNA. Independent-Samples T Test was used to test the significance. Asterisks indicate significant differences in relative expression between RNAi lines and control. (*, $p < 0.05$). *FAD3A*, *FAD3B* and *FAD3C*, cDNA vectors. pMU103, control hpRNA vector with no IR. 35S-173 and GlyP-173, control atasiRNA vectors containing only miR173 expression cassette. pMUFAD, hpRNA-producing vector targeting *GmFAD3*. 35S-FAD, GlyP-FAD, atasiRNA-generating vector containing both miR173 and modified *TAS1a* expression cassette targeting *GmFAD3*. GV3101 infiltrated and WT *N. benthamiana* leaf samples were used as negative control. (A) to (C) Normalized cDNA level of *GmFAD3A*, *GmFAD3B*, and *GmFAD3C*, respectively. The values of *FAD3A*, *FAD3B* and *FAD3C* were arbitrarily fixed to 1.0.

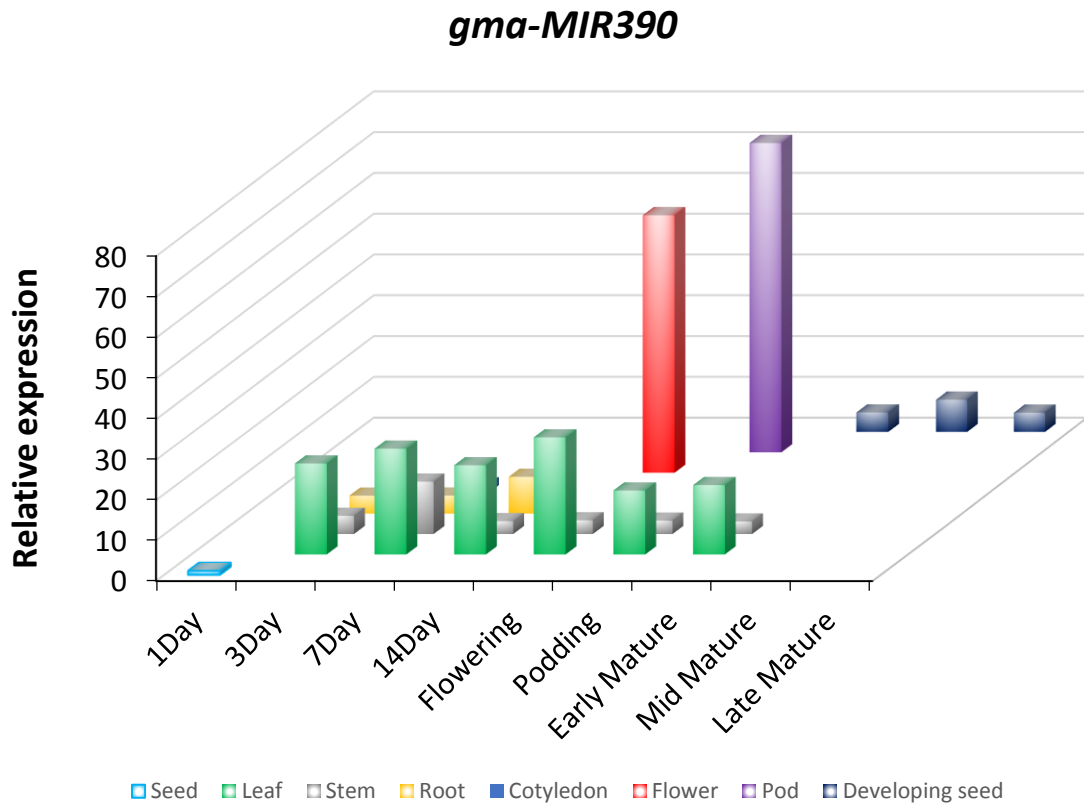


Figure 3.3 Temporal and spacial expression analysis of *gma-MIR390*. X-axis indicates soybean development stages. Bar heights represent mean relative expression level from 3 biological replications in different tissue samples. The expression level of miR390 in one day germinating seed was arbitrarily fixed to 1.0.

TABLES

Table 3.1 Sequences of atasiRNA candidates generated from the 318-nt IR.

	Position on target	Probability of being efficient siRNA	siRNA Sequence(5'->3')
1	69	0.912286	UUUCUGAUGAUGACCAUGGUG
2	149	0.905375	UUGAUCCAACCAUAGUCACGA
3	105	0.904108	UAAAUAACUCCAUUCUUGCC
4	101	0.902297	UACUCCAUUCUUGCCGCGA
5	151	0.895925	UAUUGAUCCAACCAUAGUCAC
6	243	0.891511	UUUUGCUGCUUGUGUCGCUUC
7	41	0.889758	UAUGUGACAAAGUCCAGCCAC
8	106	0.870204	UAAAUAACUCCAUUCUUGC
9	104	0.869841	AAAUAACUCCAUUCUUGCCG
10	152	0.858755	UUAUUGAUCCAACCAUAGUCA
11	246	0.855118	UGAUUUUGCUGCUUGUGUCGC
12	68	0.853079	UUCUGAUGAUGACCAUGGUGA
13	103	0.845048	AAUAACUCCAUUCUUGCCGC
14	255	0.844936	UCCAAGAACUGAUUUUGCUGC
15	20	0.84468	AUAACAAAUAUCCAUAUGGA
16	215	0.837463	UGAUAAUGAGGAAUUGAGGG
17	47	0.825638	UGCAAGUAUGUGACAAAGUCC
18	193	0.822864	AAAGAUGGUGAAUACAUGGG
19	48	0.812863	AUGCAAGUAUGUGACAAAGUC
20	192	0.807478	AAGAUGGUGAAUACAUGGGU
21	204	0.797656	AAUUUGAGGGAAAAGAUGGUG
22	142	0.797124	AACCAUAGUCACGAUCCACAG
23	102	0.787651	AUAACUCCAUUCUUGCCGCG
24	277	0.786095	UUUCUGGCUCACGGUAAUACU
25	252	0.782637	AAGAACUGAUUUUGCUGCUUG
26	138	0.779683	AUAGUCACGAUCCACAGUUGU
27	150	0.778212	AUUGAUCCAACCAUAGUCACG
28	194	0.766777	AAAAGAUGGUGAAUACAUGG
29	162	0.76444	AUGGUGAAUGUUAUUGAUCCA
30	81	0.751588	AUACCAAGGCAGUUUCUGAUG
31	146	0.751173	AUCCAACCAUAGUCACGAUCC
32	14	0.748847	AAUAUCCAUAUGGAAUCCA
33	167	0.748479	AUGUCAUGGUGAAUGUUAUUG
34	54	0.741574	AUGGUGAUGCAAGUAUGUGAC
35	249	0.739477	AACUGAUUUUGCUGCUUGUGU
36	155	0.737448	AUGUUAUUGAUCCAACCAUAG
37	116	0.735764	AGACCACCUCUAAAUAACUC
38	44	0.735201	AAGUAUGUGACAAAGUCCAGC

39	189	0.734806	AUGGUGAAUACAUGGGUGCC
40	203	0.722761	AUUUGAGGGAAAAGAUGGUGA
41	76	0.721614	AAGGCAGUUUCUGAUGAUGAC
42	294	0.719859	UGGUA AUGGUGCAGAUCUUUC
43	62	0.711927	UGAUGACCAUGGUGAUGCAAG
44	15	0.709613	AAUAUCCAAUAUGGAAUUC
45	118	0.709417	UGAGACCACCUCUAAAUAAC
46	60	0.69822	AUGACCAUGGUGAUGCAAGUA
47	50	0.681862	UGAUGCAAGUAUGUGACAAAG
48	154	0.681801	UGUUAUUGAUCCAACCAUAGU
49	297	0.679695	AAAUGGUA AUGGUGCAGAUCU
50	7	0.676354	AAUAUGGAAUCCAUAAGAGCU
51	182	0.673615	AUAACAUGGGUGCCAAUGUCA
52	183	0.662516	AAUAACAUGGGUGCCAAUGUC
53	267	0.661053	ACGGUAAUACUCUCCAAGAAC
54	259	0.659138	ACUCUCCAAGAACUGAUUUUG
55	43	0.648106	AGUAUGUGACAAAGUCCAGCC
56	129	0.637282	AUCCACAGUUGUGAGACCACC
57	191	0.636849	AGAUGGUGAAUACAUGGGUG
58	107	0.615882	CUUAAAUAACUCCAUAUCCUUG
59	156	0.598292	AAUGUUAUUGAUCCAACCAUA
60	71	0.590162	AGUUUCUGAUGAUGACCAUGG
61	180	0.589798	ACAUGGGUGCCAAUGUCAUG
62	97	0.57064	UCCAUAUCCUUGCCGCGAUACC
63	248	0.569816	ACUGAUUUUGCUGCUUGUGUC
64	40	0.569711	AUGUGACAAAGUCCAGCCACA
65	115	0.549348	GACCACCUCUAAAUAACUCC
66	185	0.543143	UGAAUAACAUGGGUGCCAAUG
67	210	0.5	AUGAGGAAUUUGAGGGAAAAG
68	132	0.5	ACGAUCCACAGUUGUGAGACC
69	32	0.491879	AAGUCCAGCCACAUAACAAAU
70	209	0.477479	UGAGGAAUUUGAGGGAAAAGA
71	278	0.474752	CUUUCUGGCUCACGGUAAUAC
72	22	0.455902	ACAUAAACAAUAUCCAUAUG
73	21	0.452734	CAUAACAAUAUCCAUAUUGG
74	251	0.424382	AGAACUGAUUUUGCUGCUUGU
75	207	0.394758	AGGAAUUUGAGGGAAAAGAUG
76	218	0.350684	AGGUGAUAUUGAGGAAUUUGA
77	198	0.297146	AGGGAAAAGAUGGUGAAUAAC
78	2	0.22725	GGAAUCCAUAAGAGCUUGAGC

(Table 3.1 continued)

* All listed atasiRNA candidates were designed by OligoWalk (http://rna.urmc.rochester.edu/cgi-bin/server_exe/oligowalk/oligowalk_form.cgi).

** atasiRNA candidate used to construct the atasiRNA construct is highlighted in red.

*** atasiRNA candidates cover the two predominant cleavage sites identified by 5' RACE are highlighted in blue.

Table 3.2 Sequences of primers used for the screen of transgenic plants.

Gene	Primer name	Primer sequence
<i>TAS1a</i>	TAS1a-F	ATAAACCTAAACCCCTAAGCGGC
	TAS1a-R	AACAGAGAGGGGCGACGGGA
<i>MIR173</i>	miR173-F	CGGCGGTCTCATCGTAATCT
	miR173-R	CACAGAGAATCACAGAGGAAAGAGAC
<i>bar</i>	Bar-F	CACCATCGTCAACCACTACATCG
	Bar-R	CAGCAGGTGGGTGTAGAGCGT

Table 3.3 Sequences of primers used for qRT-PCR analysis.

Gene	Primer name	Primer sequence
<i>FAD3A</i>	qRT- <i>FAD3A</i> -F	AGCGACACAAGCAGCAAAAT
	qRT- <i>FAD3A</i> -R	GTCTCGGTGCGAGTGAAGGT
<i>FAD3B</i>	qRT- <i>FAD3B</i> -F	CCCACCCAGTGAGAGAAAA
	qRT- <i>FAD3B</i> -R	AGCACTAGAAGTGGACTAGTTATGAAT
<i>FAD3C</i>	qRT- <i>FAD3C</i> -F	CTCAGAAATCTGGGCCATTG
	qRT- <i>FAD3C</i> -R	TCGCTAACGAAGTGATCCTGA
<i>L23</i>	qRT- <i>L23</i> -F	AAGGATGCCGTGAAGAAGATGT
	qRT- <i>L23</i> -R	GCATCGTAGTCAGGAGTCAACC

Table 3.4 Sequences of primers used in stem-loop qRT PCR.

Gene	Primer name	Primer sequence
Universal	qRT-Universal-R	GTGCAGGGTCCGAGGT
miR390	RT-miR390	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACGGCGCT
	qRT-miR390-F	GCGGCAAGCTCAGGAGGGAT
5.8S rRNA	RT-5.8S	GTTGGCTCTGGTGCAGGGTCCGAGGTATTCGCACCAGAGCCAACGACACC
	qRT-5.8S-F	TTCATTAGGGCACGCCTGCCTG

UPL probe #21 binding site is highlighted in red.

3.5 Summary of *Agrobacterium*-mediated transformation*.

Plasmid ¹	#explant	%Reg	%GD	%Green	#shoot	#GH	#LP-R	#LP-S	Bar	GOI	T1
35S-FAD	893	83	20.6	66.7	28	11	9	2	9	9	2
GlyP-FAD	1568	86.89	23.65	73.86	71	26	11	15	3	3	1
35S-173	397	95.98	30.24	84.37	13	7	4	3	3	3	0
GlyP-173	769	95.74	28.49	89.28	30	21	0	21	0	0	0

* This part of work was completed by Sha Lu, Liwen Zhou, Sandra Valdes and Muruganantham Mookkan.

¹ 35S-173 and GlyP-173 are control vectors with miR173 expression cassette under the control of 35S and Glycinin promoter, respectively. 35S-FAD and GlyP-FAD are atasiRNA vectors containing both miR173 expression cassette and modified *TAS1a* sequence targeting *GmFAD3*.

Table 3.6 Fatty acid analysis of T₂ hemizygous seeds from atasiRNAi lines.

Plasmid	Line name	Fatty acid content (%)				
		16:00	18:00	18:01	18:02	18:03
	Maverick ¹	11.8±0.5 ³	3.5±0.2	20.9±3.3	55.3±2.8ab ⁴	8.6±0.7bc
35S-FAD	L11-4-6 ²	11.5±0.4	3.9±0.3	16.7±2.1	60.2±1.4c	7.7±0.7a
	L11-4-8	11.4±0.4	3.7±0.3	16.9±1.9	60.2±1.4c	7.9±0.5ab
	L11-4-11	11.7±0.3	3.6±0.3	16.6±1.0	60.0±1.2c	8.1±0.9ab
	L11-18-1	11.2±0.3	3.5±0.2	17.1±1.3	60.5±1.1c	7.8±0.5ab
	L11-18-2	11.1±0.5	3.3±0.3	18.2±5.8	59.4±4.8c	8.1±0.8ab
	L11-18-5	11.5±0.5	3.7±0.3	17.1±1.7	59.5±0.7c	8.2±0.9ab
GlyP-FAD	ZHOU-LS-1-5-1	12±0.4	3.3±0.2	20.4±3.0	54.8±2.3a	9.5±0.7d
	ZHOU-LS-1-5-2	11.7±0.3	3.3±0.2	19.1±2.4	55.5±2.2ab	10.4±1.2e
	ZHOU-LS-1-5-3	12±0.4	3.6±0.3	18.2±1.7	57.2±1.9b	9.0±0.8cd

¹ Maverick is the WT control line with normal fatty acid content.

² Three individual plants were sampled for each RNAi line.

³ Mean fatty acid content is based on 10 T₂ seed samples for each plant.

⁴ Means within the same column followed by the same letter were not significantly different from each other at a = 0.05 level as detected by Duncan's Multiple Range Test.

LITERATURE CITED

- Adenot X, Elmayan T, Lauressergues D, Boutet S, Bouché N, Gasciolli V, Vaucheret H (2006) DRB4-Dependent *TAS3* *trans*-Acting siRNAs Control Leaf Morphology through AGO7. *Curr Biol* 16 (9):927-932
- Allen E, Xie Z, Gustafson AM, Carrington JC (2005) microRNA-Directed Phasing during *Trans*-Acting siRNA Biogenesis in Plants. *Cell* 121 (2):207-221
- Barampuram S, Zhang ZJ (2011) Recent advances in plant transformation. *Methods In Molecular Biology* (Clifton, NJ) 701:1-35. doi:10.1007/978-1-61737-957-4_1
- Beuselinck P, Sleper D, Bilyeu K (2006) An assessment of phenotype selection for linolenic acid using genetic markers. *Crop Sci* 46 (2):747-750
- Bilyeu K, Gillman JD, LeRoy AR (2011) Novel Mutant Allele Combinations Produce Soybeans Containing 1% Linolenic Acid in the Seed Oil. *Crop Sci* 51 (1):259. doi:10.2135/cropsci2010.01.0044
- Bilyeu K, Palavalli L, Sleper D, Beuselinck P (2003) Three microsomal omega-3 fatty-acid desaturase genes contribute to soybean linolenic acid levels. *Crop Sci* 43 (5):1833-1838
- Bilyeu K, Palavalli L, Sleper D, Beuselinck P (2005) Mutations in soybean microsomal omega-3 fatty acid desaturase genes reduce linolenic acid concentration in soybean seeds. *Crop Sci* 45 (5):1830-1836
- Bilyeu K, Palavalli L, Sleper DA, Beuselinck P (2006) Molecular genetic resources for development of 1% linolenic acid soybeans. *Crop Sci* 46 (5):1913-1918
- Carbonell A, Takeda A, Fahlgren N, Johnson SC, Cuperus JT, Carrington JC (2014) New Generation of Artificial MicroRNA and Synthetic *Trans*-Acting Small Interfering RNA Vectors for Efficient Gene Silencing in Arabidopsis. *Plant Physiol* 165 (1):15-29

- Chen C, Ridzon DA, Broomer AJ, Zhou Z, Lee DH, Nguyen JT, Barbisin M, Xu NL, Mahuvakar VR, Andersen MR (2005) Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res* 33 (20):e179-e179
- Chen H-M, Chen L-T, Patel K, Li Y-H, Baulcombe DC, Wu S-H (2010) 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proceedings of the National Academy of Sciences* 107 (34):15269-15274
- Chen H, Nelson R, Sherwood J (1994) Enhanced recovery of transformants of *Agrobacterium tumefaciens* after freeze-thaw transformation and drug selection. *BioTechniques* 16 (4):664-668, 670
- de Felippes FF, Ott F, Weigel D (2011) Comparative analysis of non-autonomous effects of tasiRNAs and miRNAs in *Arabidopsis thaliana*. *Nucleic Acids Res* 39 (7):2880-2889
- de la Luz Gutiérrez-Nava M, Aukerman MJ, Sakai H, Tingey SV, Williams RW (2008) Artificial trans-acting siRNAs confer consistent and effective gene silencing. *Plant Physiol* 147 (2):543-551
- Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song K, Pikaard CS (2006) Gateway - compatible vectors for plant functional genomics and proteomics. *The Plant Journal* 45 (4):616-629
- Felippes FF, Weigel D (2009) Triggering the formation of tasiRNAs in *Arabidopsis thaliana*: the role of microRNA miR173. *EMBO reports* 10 (3):264-270
- Feng L, Duan CG, Guo HS (2013) Inhibition of in vivo Slicer activity of Argonaute protein 1 by the viral 2b protein independent of its dsRNA - binding function. *Mol Plant Pathol* 14 (6):617-622
- Flores T, Karpova O, Su X, Zeng P, Bilyeu K, Sleper DA, Nguyen HT, Zhang ZJ (2008a) Silencing of GmFAD3 gene by siRNA leads to low [α]-linolenic acids (18:3) of fad3-mutant phenotype in soybean [*Glycine max* (Merr.)]. *Transgenic Research* 17 (5):839. doi:10.1007/s11248-008-9167-6
- Flores T, Karpova O, Su X, Zeng P, Bilyeu K, Sleper DA, Nguyen HT, Zhang ZJ (2008b) Silencing of GmFAD3 gene by siRNA leads to low α -linolenic acids (18: 3) of

fad3-mutant phenotype in soybean [*Glycine max* (Merr.)]. *Transgenic Res* 17 (5):839-850

Garcia D A miRacle in plant development: role of microRNAs in cell differentiation and patterning. In: *Seminars in cell & developmental biology*, 2008. vol 6. Elsevier, pp 586-595

Garcia D, Collier SA, Byrne ME, Martienssen RA (2006) Specification of Leaf Polarity in *Arabidopsis* via the *trans*-Acting siRNA Pathway. *Curr Biol* 16 (9):933-938

Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, Kasschau KD, Carrington JC (2007) Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in *Arabidopsis* reveals dependency on miRNA-and tasiRNA-directed targeting. *The Plant Cell Online* 19 (3):926-942

Hunter C, Willmann MR, Wu G, Yoshikawa M, de la Luz Gutiérrez-Nava M, Poethig SR (2006) Trans-acting siRNA-mediated repression of ETTIN and ARF4 regulates heteroblasty in *Arabidopsis*. *Development* 133 (15):2973-2981

Joshi T, Yan Z, Libault M, Jeong D-H, Park S, Green PJ, Sherrier DJ, Farmer A, May G, Meyers BC (2010) Prediction of novel miRNAs and associated target genes in *Glycine max*. *BMC Bioinformatics* 11 (Suppl 1):S14

Li J-F, Chung HS, Niu Y, Bush J, McCormack M, Sheen J (2013) Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. *The Plant Cell Online* 25 (5):1507-1522

Liu D, Shi L, Han C, Yu J, Li D, Zhang Y (2012) Validation of reference genes for gene expression studies in virus-infected *Nicotiana benthamiana* using quantitative real-time PCR. *PloS one* 7 (9):e46451

Lu ZJ, Mathews DH (2008) OligoWalk: an online siRNA design tool utilizing hybridization thermodynamics. *Nucleic Acids Res* 36 (suppl 2):W104-W108

Marin E, Jouannet V, Herz A, Lokerse AS, Weijers D, Vaucheret H, Nussaume L, Crespi MD, Maizel A (2010) miR390, *Arabidopsis* TAS3 tasiRNAs, and their AUXIN RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *The Plant Cell Online* 22 (4):1104-1117

- Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC (2008) Specificity of ARGONAUTE7-miR390 Interaction and Dual Functionality in *TAS3* Trans-Acting siRNA Formation. *Cell* 133 (1):128-141
- Ossowski S, Schwab R, Weigel D (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. *The Plant Journal* 53 (4):674-690
- Peragine A, Yoshikawa M, Wu G, Albrecht HL, Poethig RS (2004) SGS3 and SGS2/SDE1/RDR6 are required for juvenile development and the production of trans-acting siRNAs in Arabidopsis. *Genes Dev* 18 (19):2368-2379
- Radwan O, Liu Y, Clough SJ (2011) Transcriptional analysis of soybean root response to *Fusarium virguliforme*, the causal agent of sudden death syndrome. *Mol Plant-Microbe Interact* 24 (8):958-972
- Rajagopalan R, Vaucheret H, Trejo J, Bartel DP (2006) A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes Dev* 20 (24):3407-3425
- Rhee Y, Staswick PE (1992) Nucleotide Sequence of a Soybean Vegetative Storage Protein vspB Gene. *Plant Physiology* 98 (2):794-795
- Subramanian S, Fu Y, Sunkar R, Barbazuk WB, Zhu J-K, Yu O (2008) Novel and nodulation-regulated microRNAs in soybean roots. *BMC Genomics* 9 (1):160
- Turner M, Yu O, Subramanian S (2012) Genome organization and characteristics of soybean microRNAs. *BMC Genomics* 13 (1):169
- Varkonyi-Gasic E, Wu R, Wood M, Walton EF, Hellens RP (2007) Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant methods* 3 (1):12
- Vaucheret H (2005) MicroRNA-dependent trans-acting siRNA production. *Science Signaling* 2005 (300):pe43

- Vazquez F, Vaucheret H, Rajagopalan R, Lepers C, Gascioli V, Mallory AC, Hilbert J-L, Bartel DP, Crété P (2004) Endogenous *trans*-Acting siRNAs Regulate the Accumulation of *Arabidopsis* mRNAs. *Mol Cell* 16 (1):69-79
- Williams L, Carles CC, Osmont KS, Fletcher JC (2005) A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the *Arabidopsis* ARF2, ARF3, and ARF4 genes. *Proceedings of the National Academy of Sciences of the United States of America* 102 (27):9703-9708
- Wright TR, Shan G, Waish TA, Lira JM, Cui C, Song P, Zhuang M, Arnold NL, Lin G, Yau K, Russell SM, Cicchillo RM, Peterson MA, Simpson DM, Zhou N, Ponsamuel J, Zhang Z (2010) Robust crop resistance to broadleaf and grass herbicides provided by aryloxyalkanoate dioxygenase transgenes. *Proceedings of the National Academy of Sciences of the United States of America* 107 (47):20240-20245. doi:10.1073/pnas.1013154107
- Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC (2005a) Expression of *Arabidopsis* MIRNA genes. *Plant Physiol* 138 (4):2145-2154
- Xie Z, Allen E, Wilken A, Carrington JC (2005b) DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* 102 (36):12984-12989
- Yoshikawa M, Peragine A, Park MY, Poethig RS (2005) A pathway for the biogenesis of trans-acting siRNAs in *Arabidopsis*. *Genes Dev* 19 (18):2164-2175
- Zeng P, Polacco JC, Zhang Z, Vадnais DA (2004) Refined glufosinate selection in *Agrobacterium*-mediated transformation of soybean (*Glycine max* (L.) Merrill). *Plant Cell Reports* 22 (7):478-482
- Zhang Z, Clemente TE, Staswick P, Xing A (1999a) The use of glufosinate as a selective agent in *Agrobacterium*-mediated transformation of soybean. *Plant Cell, Tissue and Organ Culture* 56 (1):37-46
- Zhang Z, Xing A, Staswick P, Clemente TE (1999b) The use of glufosinate as a selective agent in *Agrobacterium*-mediated transformation of soybean. *Plant Cell Tiss Org Cult* 56 (1):37-46

Zhang ZJ (2014) Artificial trans-acting small interfering RNA: a tool for plant biology study and crop improvements. *Planta* 239 (6):1139-1146

VITA

Sha Lu was born January 8, 1985 in Jiaxing, China. She graduated from Tongxiang High School in 2003 and study biology at Zhejiang University. Sha earned her bachelors of science in 2008 and was awarded Chinese Government Scholarship to study for a Ph.D degree at the University of Missouri. She started her Ph.D work with Dr. Zhanyuan Zhang in the Division of Plant Science since 2008.