Plant Biotechnology Journal Open Access

Plant Biotechnology Journal (2022) **20**, pp. 1069–1084 doi: 10.1111/pbi.13786

Artificial microRNA guide strand selection from duplexes with no mismatches shows a purine-rich preference for virus- and non-virus-based expression vectors in plants

Yen-Wen Kuo and Bryce W. Falk*

Department of Plant Pathology, University of California Davis, Davis, CA, USA

Received 8 August 2020; revised 18 January 2022; accepted 24 January 2022.
*Correspondence: (Tel +1-530-752-0302; fax +5307525674; email: bwfalk@ucdavis.edu)

Keywords: artificial microRNA, guide strand selection, virus-based expression vector, small RNA deep sequencing.

Summary

Artificial microRNA (amiRNA) technology has allowed researchers to direct efficient silencing of specific transcripts using as few as 21 nucleotides (nt). However, not all the artificially designed amiRNA constructs result in selection of the intended ~21-nt guide strand amiRNA. Selection of the miRNA guide strand from the mature miRNA duplex has been studied in detail in human and insect systems, but not so much for plants. Here, we compared a nuclear-replicating DNA viral vector (tomato mottle virus, ToMoV, based), a cytoplasmic-replicating RNA viral vector (tobacco mosaic virus, TMV, based), and a non-viral binary vector to express amiRNAs in plants. We then used deep sequencing and mutational analysis and show that when the structural factors caused by base mismatches in the mature amiRNA duplex were excluded, the nucleotide composition of the mature amiRNA region determined the guide strand selection. We found that the strand with excess purines was preferentially selected as the guide strand and the artificial miRNAs that had no mismatches in the amiRNA duplex were predominantly loaded into AGO2 instead of loading into AGO1 like the majority of the plant endogenous miRNAs. By performing assays for target effects, we also showed that only when the intended strand was selected as the guide strand and showed AGO loading, the amiRNA could provide the expected RNAi effects. Thus, by removing mismatches in the mature amiRNA duplex and designing the intended guide strand to contain excess purines provide better control of the guide strand selection of amiRNAs for functional RNAi effects.

Introduction

Since their discovery, microRNAs (miRNAs) have been evaluated for a broad range of applications. Artificial microRNAs (amiRNAs) have been used for applications including mediating virus resistance in plants (Carbonell *et al*., 2016; Niu *et al*., 2006; Wagaba *et al*., 2016) and animals (Hutcheson *et al*., 2015; Motavaf *et al*., 2014), and as potential biomarkers for diseases (Xu *et al*., 2016; Zhao *et al*., 2016). miRNAs are processed from endogenously derived non-coding RNAs and are important regulators of gene expression. In plants, primary miRNAs (pri-miRNAs) show specific secondary structures and are first processed into precursor miRNAs (pre-miRNAs) and then processed into miRNA duplexes in the nucleus by DCL1. The miRNA duplexes are then relocated to cytoplasm. Subsequently, miRNA duplexes load into argonaute (AGO) proteins in the RISC complex (Rogers and Chen, 2013; Zhang *et al*., 2015), and one of the strands is selected to become the guide strand (Meijer *et al*., 2014; Schwarz *et al*., 2003). When the miRNA incorporated RISC encounters the targeted RNA, the latter is then degraded or translationally suppressed (Reis *et al*., 2015). The miR/miR* (the guide strand/passenger strand) duplexes in both plants and animals are similar with ~20- to 24-nt in length and imperfect complementarity between the two strands and with 2nt 3′-overhangs (Bartel, 2004; Voinnet, 2009).

The amiRNA technology for plant applications has offered alternative and improved approaches over the existing gene-silencing approaches (mainly long dsRNA or hairpin RNA precursor-induced siRNAs) and uses endogenous pri-miRNA backbones but replaces the miRNA and miRNA* with amiRNA/amiRNA* sequences that are artificially designed to target transcripts of interest. One of the strategies for generating amiRNAs is by replacing the mature miRNA sequences of the pri-miRNA backbone with the sequence of intended targets without mismatches between amiRNA/amiRNA* (Liang *et al*., 2012; Niu *et al*., 2006). While many amiRNA approaches have been applied through producing transgenic plants, viruses have also been used as alternative amiRNA delivery methods (Tang *et al*., 2010). Viruses offer advantages as often it is easier to engineer a virus than the host, and because viruses replicate in their hosts, they can generate high amount of amiRNAs. When artificially expressing miRNAs with pri-miRNA backbones in animal systems using viral vectors, most of the selected viruses were nuclear-replicating viruses (Herrera-Carrillo *et al*., 2017; Honda *et al*., 2016). In plants, it has been reported that a cytoplasmic-replicating virus, barley stripe mosaic virus (BSMV, a virgavirus, a (+) ssRNA virus) was able to express amiRNAs initiated by a pri-miRNA and a pre-miRNA backbone (Jian *et al*., 2017), which is somewhat unexpected. However, if plant cytoplasmic-replicating viruses could be used to express specific amiRNAs in plants, they

offer significant advantages over more traditional transgenic approaches. Therefore, to study this further, we included another cytoplasmic-replicating virgavirus, along with a nuclear-replicating DNA virus and a non-viral vector to express amiRNAs from a pri-miRNA backbone and compared the small RNA patterns and the quality and quantity of amiRNAs produced from each vector.

In addition to the vectors used to express amiRNAs, other factors can affect the effectiveness of amiRNAs, and not all designed amiRNAs perform as intended in plants. Although there are computer programs to optimize design of mature amiRNA sequences such as WMD3 amiRNA designer and P-SAMS (Fahlgren *et al*., 2016), some of these amiRNAs still fail to deliver the desired target effects (Arroyo *et al*., 2014; Carbonell *et al*., 2019). Although some factors of guide strand selection were included when designing amiRNAs (Arroyo *et al*., 2014; Fellmann *et al*., 2011), additional factors may be omitted in the designs. Therefore, understanding factors determining miRNA or amiRNA biogenesis and guide strand selection will help researchers have more control over designing efficient functional amiRNAs for use in plants.

Studies of human and insect miRNAs showed that both strands of the same miRNA duplexes (miR-5p and -3p arms) can be selected to become the guide strand under different conditions (Li *et al*., 2012a,b). This is critical to consider for attempting to use amiRNAs in biotechnology as different miRNA-directed effects can result depending on which strand is selected (Griffiths-Jones *et al*., 2011; Ohanian *et al*., 2013; Rubio *et al*., 2013; Yang *et al*., 2013). Several factors can be involved in strand selection, including miRNA processing-related proteins, post-transcriptional modifications, thermostability of miRNA duplexes and 5′-end nucleotides of the strands. Furthermore, it has been suggested that the hydrophobic interactions between purines and the aromatic residues in the AGO PAZ domain could result in the purine-rich strand becoming the guide strand leading to the preferential loading of the duplex in a specific orientation for vertebrates and flies (Hu *et al*., 2009; Meijer *et al*., 2014). Therefore, not only the structure caused by the mismatches in the miRNA duplex, but also the nucleotide sequence/composition affected the strand selection.

Based on our initial next generation sequencing data, we found that the guide strand of our amiRNAs could originate from the 5p or 3p strand, even when using the same pri-miRNA backbone. We found that the strand with excess purines was preferentially selected as the guide strand. Therefore, we tested this further by constructing additional amiRNA constructs with different numbers of purine residues in each strand and assessed amiRNA guide strand selection. We developed a plant virus-based system to express high amounts of specific amiRNAs in plants and showed desired target effects.

Results

Cloning and expression of amiRNAs in *Nicotiana benthamiana* plants

Plant pri-miR319a was used as the backbone for producing amiRNAs in *N. benthamiana* plants. Both arms of the mature miRNA region of the pri-miR319a were replaced to produce amiRNAs. Initially, two amiRNAs, amiRA2 and amiRA2c, were designed based on using the online program WMD3 (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi) (Ossowski *et al*., 2008) such that amiRA2 5p contained reverse complementary sequences compared with the 5p strand of amiRA2c (Figure 1).

The pri-miR319a backbones containing the amiRNA sequences were cloned into three different vectors for transient expression assays in plants. In our studies, we evaluated using a nuclear-replicating tomato mottle virus DNA A component viral vector (TAV) (Hou and Gilbertson, 1996), a cytoplasmic-replicating RNA viral vector derived from tobacco mosaic virus (TRBO) (Lindbo, 2007) and a non-viral binary plasmid vector (pGWB2) to express amiRNAs. The viruses can autonomously replicate in plant host cells and might give higher levels of the amiRNAs than would the standard binary plasmid, and they could also be delivered more easily than by producing transgenic plants, therefore, broadening the potential for amiRNA applications in plants. The resulting plasmids were transformed into *Agrobacterium tumefaciens* and infiltrated into leaves of *N. benthamiana* plants (Figure 2a). Northern blot hybridization and PCR analyses were performed to confirm the replication of recombinant TRBO and TAV viral vectors respectively (Figure 2b,c). A TAV common region (CR) mutant (5-nucleotide mutation) was used as a negative control for TAV replication. Northern blot hybridization analysis showed strong TRBO replication, subgenomic mRNAs and full-length negative genomic strand RNA accumulation. In addition, the small subgenomic RNA containing the pri-amiRNA coding sequence was abundant (Figure 2b, lower band). PCR analysis showed that both of the TAV constructs (harbouring amiRA2 and one with amiRA2C) replicated, while no PCR product was seen for the non-infectious TAV CR mutant (Figure 2c).

Small RNA northern blot analyses suggested differential amiRNA production

To determine whether the intended amiRNAs were generated in these plants, total RNAs extracted from the infiltrated tissues were used for small RNA northern blot hybridization analyses (Figure 2d). The amiRNA accumulation levels were analyzed by using probes designed to hybridize with amiRA2-5p and amiRA2c-5p. The results showed qualitative and quantitative differences in accumulation levels for each (Figure 2d). The amiRA2c-5p showed much higher accumulation levels for all three vectors, compared with the amiRA2-5p (Figure 2d). Furthermore, for the virus-driven expression, more than one size of small RNA was detected for both amiRA2c-5p and amiRA2-5p for both viruses (Figure 2d). TAV gave a prominent 21-nt product, but also a less intense signal, suggesting a product of ~ 22-nt. The TRBO-driven amiRNA analysis showed a signal likely representing a 21-nt amiRNA; however, there were two additional signals one representing a ~24-nt small RNA and a larger product beyond the range of our markers (Figure 2d). By contrast, the non-virus-based binary plasmid pGWB2 gave a single RNA representing the desired ~ 21-nt amiRNA. Based on intensity of hybridization signals, these analyses also suggested that the TAV vector gave the highest accumulation of amiRA2c and amiRA2, compared to those expressed from TRBO or pGWB2.

The TAV vector produced high quality and quantity of amiRNAs

The above results suggested that amiRNAs were expressed using all three vectors, but qualitative and quantitative differences were apparent. Furthermore, the less intense hybridization signal intensities obtained for the amiRA2 vs. amiRA2c were unexpected. We expected that the guide strand of both amiRA2 and amiRA2c would originate from the same arm (5P or 3P) because

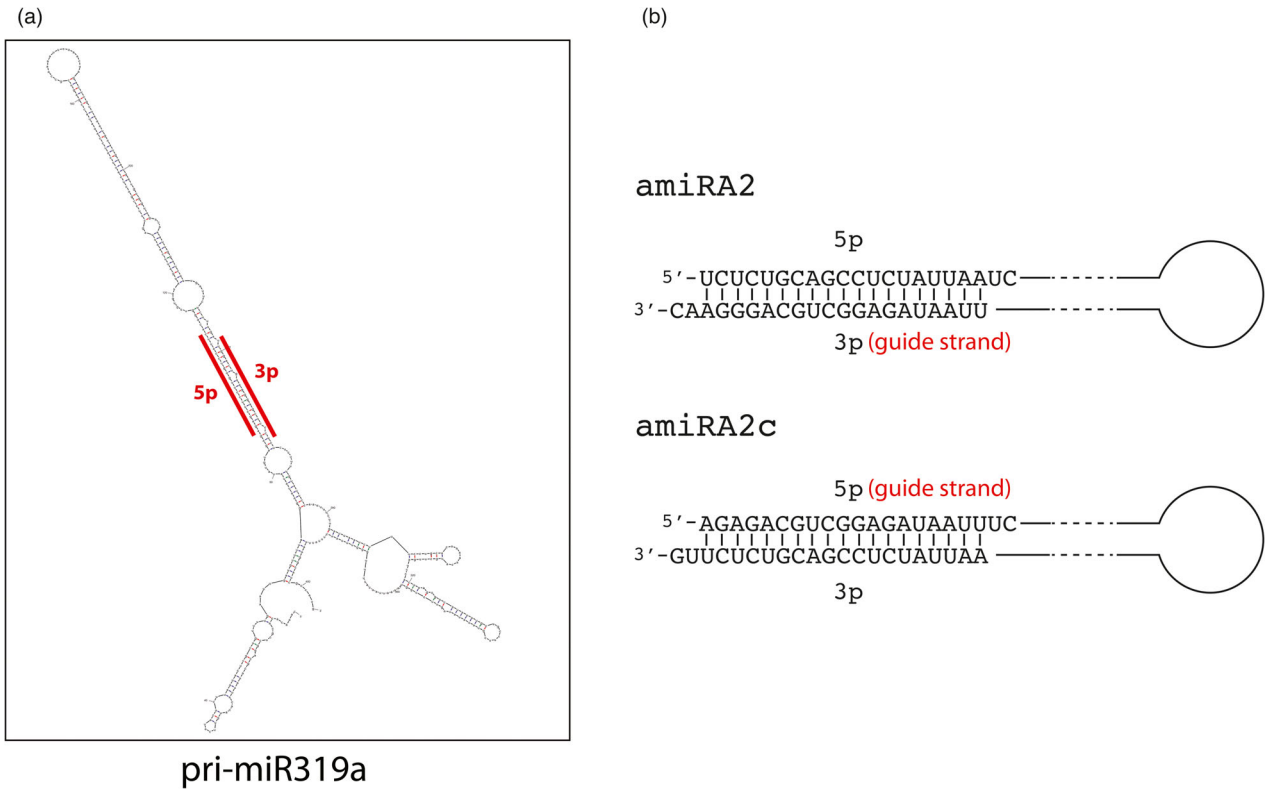


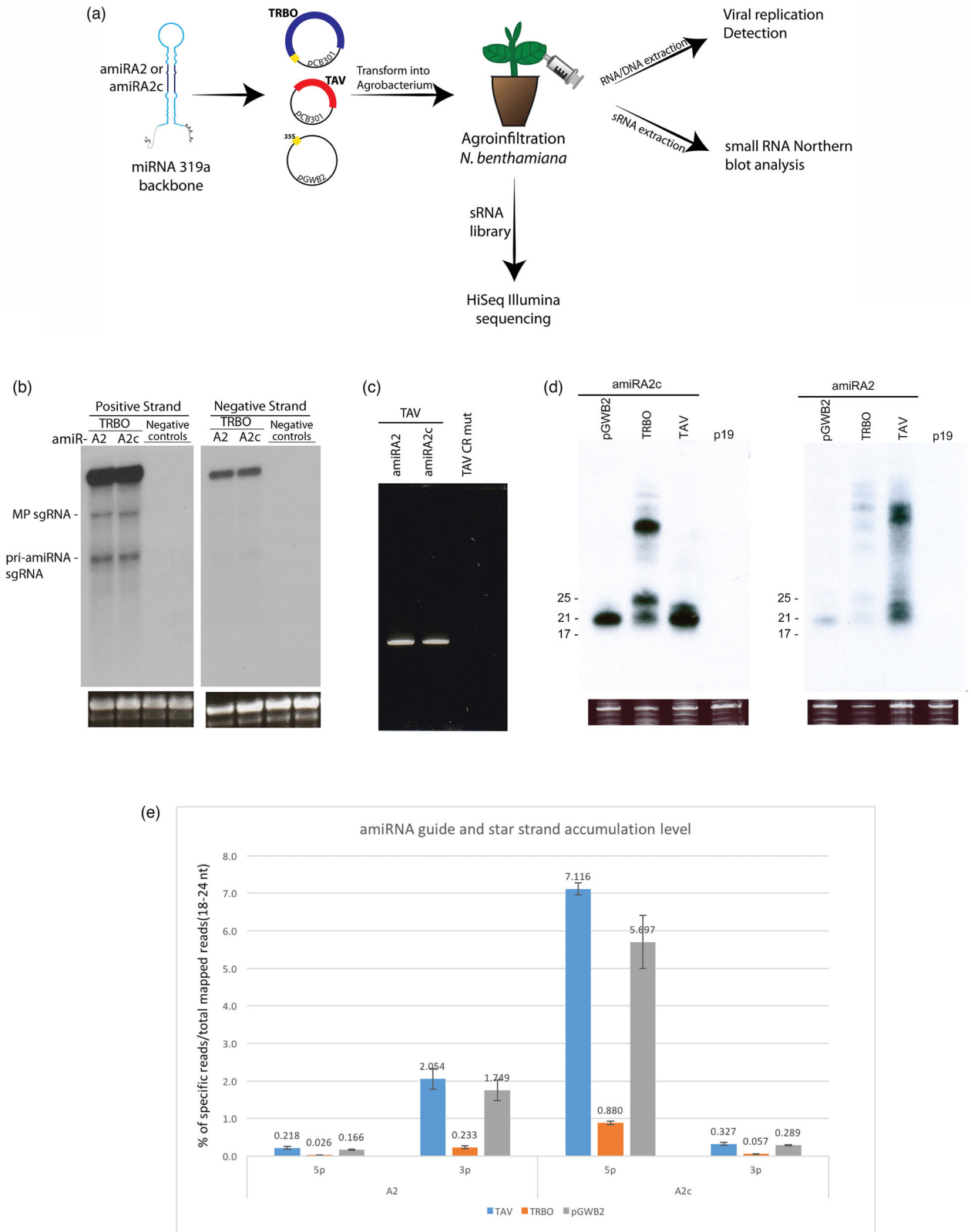
Figure 1 The predicted folding structure of pri-miR319a and partial pri-amiRNAs. (a) 5p and 3p arms of the miRNA duplex are shown in the box. The mature miR319a duplex region contains 3 mismatches between its 5p and 3p strands. (b) Sequences and the guide strands of amiRA2 and amiRA2c duplex. The mismatches between the 5p and 3p strands were removed from the amiRA2 and amiRA2c constructs. The guide strands of amiRA2 and amiRA2c that were identified from this study was labelled.

they shared the same pri-miRNA backbone and were expressed in the same cell types under the same conditions, and therefore, the resulting signal intensities should be similar. However, the results indicated that the guide strand selection for amiRA2 and amiRA2c may be different.

We further analyzed the amiRNAs by HiSeq small RNA Illumina sequencing. Three replicate samples for each vector construct were sequenced. The sequencing reads were used to map back to the precursor and expression vector. The deep sequencing analysis allowed for precise quantitative (Figure 2e) and qualitative (Figures S1–S3) analysis of amiRNAs. The small RNAs mapping to TRBO showed small RNAs mapping back to the entire TRBO sequence, suggesting that specific amiRNAs were not the most abundant small RNAs generated here. The results also showed that the ~21-nt RNAs expressed from TRBO did not have the precise cleavages as predicted for the miRNA pathway. 21-nt RNAs were most abundant, but the intended 3p and 5p amiRNAs were not more abundant than were small RNAs derived from other regions of the TRBO genomic RNA (Figure S1). By contrast, the bioinformatic analyses showed that both TAV and pGWB2 gave specific production of 21-nt amiRNAs, indicating precise miRNA cleavages in the sequence (Figure S2 and S3). TAV-derived amiRA2 (TAV-amiRA2) and TAV-amiRA2c both showed small RNAs mapping to the entire TAV sequence, but the most prominent peaks were for the amiRNA regions, particularly for TAV-amiRA2c. Similarly, pGWB2 gave small RNAs mapping across the T-DNA region, but the amiRNA regions showed very prominent peaks showing that the intended amiRNAs were specifically generated (Figure S2). Although the non-viral vector, pGWB2, expressed good quality and quantity amiRNAs, the TAV vector gave the greatest quantity of specific amiRNAs (Table 1 and Figure 2e).

Deep sequencing data revealed alternate amiRNA strand selection

Previous work has shown that when one strand was selected as the guide strand, the other typically is degraded (Warf *et al.*, 2011; Zinovyeva *et al.*, 2015); therefore, the strand showing significantly higher accumulation levels was interpreted to be the guide strand. Our analysis showed that the 3p strand of amiRA2 had accumulation levels approximately 10-fold higher than the 5p strand indicating that the amiRA2 3p strand was the guide strand (Figure 2e and Table 1). However, the accumulation levels of the 5p strand of amiRA2c were 15 to 20 times higher than its 3p strand, indicating that, unlike amiRA2, the 5p strand of amiRA2c was the guide strand (Figure 2e and Table 1). Further analysis showed that the 3p arm of amiRNA2 contained 14 nt that were either A or G (purines) while the 5p arm has only six purines. Similarly, the 5p arm of amiRA2c has 13 purines while the 3p arm has 7, and in both cases, the purine-rich strands of amiRA2 and amiRA2c were selected as the guide strands. In humans, it has been suggested that an excess of purines (A/G) could be one of the factors affecting strand selection (Hu *et al.*, 2009), but this is not known for plants.



Purine content and guide RNA selection

The above results led us to further examine whether the purine content within the miRNA duplex affected guide strand selection for plants. Therefore, additional amiRNAs, amiRA1, A1c, A3, A3c, G1, G1c, G2 and G2c with different nucleotide sequences and purine content with no mismatches in the miRNA duplex regions were designed (Table 2). Based on our previous small RNA northern blot analysis and deep sequencing verification shown above, we decided to use quantitative small RNA northern blot

Figure 2 Production of amiRNAs in plants. (a) The primary miR319a backbone was used for producing amiRNAs and cloned into three different vectors. The two viral vectors, TRBO and TAV, are in the binary vector pCB301 backbone, and the non-viral vector is the binary vector pGWB2. Both TRBO and pGWB2 are driven by 35S promoter of cauliflower mosaic virus. (b) The replication of TRBO was confirmed by northern blot hybridization. Genomic RNA, movement protein (MP) subgenomic RNA (sgRNA) and pri-amiRNA subgenomic RNA were detected from the TRBO-infiltrated tissues with the probe-detecting sense (positive strand) RNA. The negative strand of the TRBO genomic RNAs was detected with the probe (positive strand), and the negative strand is only derived during virus replication. The negative controls are the pCB301 vector carrying non-TMV sequences. (c) PCR was used to confirm the replication of TAV. TAV CR mutant, which contains a 5-nucleotide mutation in the non-coding common region of ToMoV to abolish viral replication, was used as a negative control. (d) amiRA2 and amiRA2c accumulation levels produced with the three different vectors were detected by the specific 21-nt anti-sense DNA probes by northern blot analysis. P19, the silencing suppressor alone was used as a negative control. (e) Percentages of amiRA2 and amiRA2c guide and star strand reads accumulation level in total mapped 18-25 nt reads in three different vectors listed in the Table 1 are shown in a bar chart.

analysis to compare guide strand selection for the analyses here. Comparisons of the hybridization intensities showed that in all but one case, strands with excess of purines (A/G), whether in the 5p or 3p arm, were selected to be the guide strand (Figure 3 and Table 2). The one exception was the A1c 5p strand (nine purines), which was selected over the 3p strand (11 purines), and both strands had a A as the 5′ nucleotide. However, for amiRG1, the 3p strand, which contained nine purines and a 5′ terminal A, was not selected as the guide strand, while the 5p strand, which contained 11 purines and a 5′ terminal U, was selected as the guide strand (Figure 3 and Table 2).

The artificial miRNAs loaded into AGO2 but not AGO1

To confirm that the deep sequencing and northern blot analyses reflect amiRNA strand selection, we examined whether the amiRNAs were loaded into AGO proteins. We cloned HA-tagged *Arabidopsis* AGO1 and AGO2 expression constructs for co-immunoprecipitation (co-IP) analyses. 3X-HA tag was cloned at the N′ terminus of the AGO1 sequence and 1X-HA at the N′ terminus of the AGO2 sequence. Both constructs were driven by the 35S promoter of cauliflower mosaic virus. These were cloned into a binary vector pCB301 for agroinfiltration assays.

Table 1 Specific nucleotide sequences of amiRA2 and amiRA2c-5p- and 3p-arm deep sequencing read numbers and percentages from TAV, TRBO and pGWB2 vectors. Three biological deep sequencing repeats are shown in the table

		TAV			TRBO			pGWB2		
	Read (DNA) sequence	Repeats*	Reads #/total mapped reads†	%	Repeats	Reads #/total mapped reads	%	Repeats	Reads #/total mapped reads	%
amiRA2										
5p	TCTCTGCAGCCTCTATTAATC	2A15	27,759/19,723,127	0.141	4A20	1,917/6,602,726	0.029	2C17	34,127/ 18,327,238	0.186
		4C22	38,404/17,939,282	0.214	1A13	3,228/10,222,801	0.032	3B24	56,691/ 29,248,561	0.194
		2C	33,807/11,257,031	0.300	2B	1,580/8,983,459	0.018	2D19	22,644/ 19,172,383	0.118
3p	TTAATAGAGGCTGCAGAGAAC	2A15	297,183/ 19,723,127	1.507	4A20	16,585/6,602,726	0.251	2C17	328,076/ 18,327,238	1.790
		4C22	364,142/ 17,939,282	2.030	1A13	30,400/ 10,222,801	0.297	3B24	667,313/ 29,248,561	2.282
		2C	295,525/ 11,257,031	2.625	2B	13,490/8,983,459	0.150	2D19	225,515/ 19,172,383	1.176
amiRA2c										
5p	AGAGACGTCGGAGATAATTTC	2B16	1,556,036/ 22,411,150	6.943	1B14	81,902/8,346,072	0.981	2D18	1,187,293/ 23,933,562	4.961
		3A23	2,044,288/ 27,323,697	7.482	4B21	51,571/5,972,788	0.863	3C14	921,371/ 19,183,756	4.803
		1C	2,307,081/ 33,326,226	6.923	1A	104,213/ 13,096,528	0.796	1B	2,243,122/ 30,610,608	7.328
3p	AATTATCTCCGACGTCTCTTG	2B16	86,352/22,411,150	0.385	1B14	4,977/8,346,072	0.060	2D18	58,944/ 23,933,562	0.246
		3A23	92,598/27,323,697	0.339	4B21	3,739/5,972,788	0.063	3C14	63,489/ 19,183,756	0.331
		1C	85,140/33,326,226	0.255	1A	6,286/13,096,528	0.048	1B	88,591/ 30,610,608	0.289

*The sample codes of deep sequencing repeats.

†The read number of the exact 21-nt amiRNA read sequences over the total 18- to 24-nt small RNA reads.

Table 2 The additional amiRNAs and corresponding nucleotide sequences

amiRNA	Strand	Sequence (5′–3′)	# of A/G
amiRA2	5p	UCUCUGCAGCCUCUAUUAAUC	6
	3p	**UUAAUAGAGGCUGCAGAGAAC**	**14**
amiRA2c	**5p**	**AGAGACGUCGGAGAUAAUUUC**	**13**
	3p	AAUUAUCUCCGACGUCUCUUG	7
amiRA1	5p	UGGACCCUACCAUUAAGAAUC	10
	3p	**UUCUUAAUGGUAGGGUCCAGC**	**10**
amiRA1c	5p	ACCUGGGAUGGUAAUUCUUUC	9
	3p	**AAGAAUUACCAUCCCAGGUCG**	**11**
amiRA3	**5p**	**AUGAUGCUGACAAGACAGAUC**	**13**
	3p	UCUGUCUUGUCAGCAUCAUCC	6
amiRA3c	5p	UACUACGACUGUUCUGUCUUC	6
	3p	**AGACAGAACAGUCGUAGUAGG**	**15**
amiRG1	**5p**	**UGAAUUAGAUGGUGAUGUUUC**	**11**
	3p	AACAUCACCAUCUAAUUCAAC	9
amiRG1c	5p	ACUUAAUCUACCACUACAAUC	8
	3p	**UUGUAGUGGUAGAUUAAGUUG**	**12**
amiRG2	**5p**	**UUAUGUACAGGAAAGAACUUC**	**12**
	3p	AGUUCUUUCCUGUACAUAACC	7
amiRG2c	5p	AAUACAUGUCCUUUCUUGAUC	7
	3p	**UCAAGAAAGGACAUGUAUUGG**	**14**

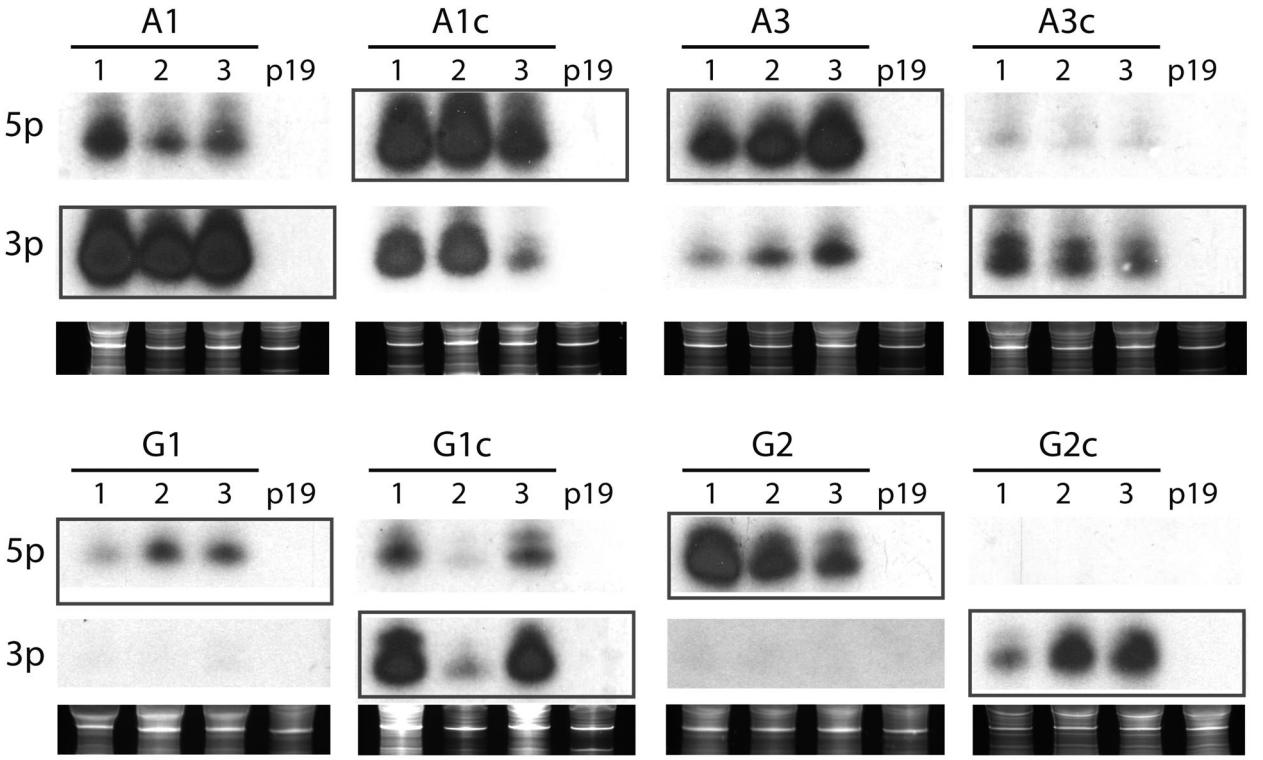
*Sequences in bold indicate the guide strand of the miRNAs.

Each *Agrobacterium tumefaciens* culture carrying TAV-amiRNA constructs were co-infiltrated into leaves of *N. benthamiana* plants along with 3X-HA-AGO1 or 1X-HA-AGO2 *Agrobacterium* cultures. The infiltrated *N. benthamiana* tissues were collected at 4 dpi and used for co-IP analyses. The Western blot analyses of the IP fraction showed that both 3X-AGO1 and 1X-HA-AGO2 were successfully pulled down (Figure 4). The total RNA of each sample was then extracted from the IP fraction. Different from the total RNA northern blot analyses, which showed an extra product (~22-nt) for TAV-amiRNAs, the co-IP RNA northern blot analyses showed that only the 21-nt amiRNAs were loaded into argonaute proteins. For most of the TAV-amiRNAs tested in our studies, only one strand of the mature TAV-amiRNA was strongly detected with the argonaute proteins with one exception of amiRA1c. Co-IP RNA northern blot analyses showed that both strands of TAV-amiRA1c were loaded into AGO2. These analyses also showed that the TAV-amiRNAs were predominantly loaded into the AGO2 but not AGO1 except the amiRA1, of which the guide strand (3p) was detected in both AGO1 and AGO2 (Figure 4b). These results showed that the amiRNAs expressed from the TAV vector were loaded into argonaute proteins.

The selected amiRNA guide strand induced intended target effects

To determine whether the intended amiRNA strand was selected and could induce desired down-stream target silencing effects and to confirm the correlation of the purine-rich factor of the guide strand selection, we first performed reporter assays in *N. benthamiana* plants. We added 35S promoter and NOS terminator in the TAV viral vector to express enhanced green fluorescent protein (EGFP) of which the target sequences of different amiRNAs were cloned at the 3′ UTR of the EGFP ORF (Figure 5a). The TAV-amiR-GFP-target clones were agroinfiltrated into *N. benthamiana* plants for EGFP expression analyses. The infiltrated tissues were used for western blot analyses, and the expression level of the EGFP was quantified by the software imageJ. We first cloned the target sequences of the 3p strands of amiRA1, amiRA2 and amiRA3 and compared the RNAi effects with the same target sequences that paired with amiRA1c, amiRA2c and amiRA3c, of which the amiRNAs would not be able to recognize the cloned target sequences. Constructs without amiRNA sequences but with the GFP-target sequences were also used as controls, which were labelled as EV. The results showed that the EGFP expression was down-regulated in the assays of amiRA1-EGFP_A1_3p target and amiRA2-EGFP_A2_3p target, while the assays of amiRA3-EGFP_A3_3p target did not show statistically significant down-regulation of EGFP expression (Figure 5). These results were in agreement with our strand selection preference and AGO protein-loading results: the guide strand of amiRA1 and amiRA2 were the 3p strand while that of the amiRA3 was the 5p strand. We then tested and compared the RNAi effects of amiRA1, amiRA1c, amiRA2, amiRA2c, amiRA3 and amiRA3c by cloning the target sequences of the 5p or 3p strand of each amiRNA. All the assays showed down-regulation of GFP protein of the corresponding target sequences to the amiRNA strands that were shown to be loaded into AGO protein (Figure 6 and S4). The representative Western blot results are presented in Figure 6, and the complete results of western blots that were used for statistics analyses were presented in supplemental Figure 4. Those results of the reporter assays showed that the strand of amiRNAs, which contained excess purines, showed AGO loading in the co-IP assays (Figure 4) and also showed RNAi targeting effects in the reporter assays (Figure 5 and 6). We next chose to target an endogenous plant RNA.

We constructed TAV-amiRNAs intended to target the phytoene desaturase (PDS) mRNA of *N. benthamiana*. Different numbers of purines were designed for the guide strands of different TAV-amiRPDS constructs. The TAV backbone of one amiRPDS construct was mutated in order to eliminate its ability to replicate and generate amiRNAs. 5p strands of all six the amiRPDS constructs, amiRPDS1, amiRPDS1_TAVmut, amiRPDS3, amiRPDS4, amiRPDS5 and amiRPDS6, were designed to be complementary to the PDS mRNA. The 5p strand of amiRPDS1, amiRPDS1_TAVmut, amiRPDS3, amiRPDS4 and amiRPDS5 were designed to be purine-rich, while the purine-rich strand of amiRPDS6 was the 3p strand, which would not target the NbPDS mRNA but the negative strand of NbPDS (Table 3). Because there is no mismatch in the amiRNA duplex, we suspected that the desired effects induced by those amiRNAs may be detectable at the RNA level. Therefore, we analyzed the relative accumulation level of NbPDS transcripts of the tested plants using cytochrome C oxidase (COX) as a reference. Our results of reverse transcription-quantitative PCR (RT-qPCR) analyses showed that amiRPDS1, amiRPDS3, amiRPDS4 and amiRPDS5 showed reduced NbPDS transcript accumulation levels, while the NbPDS transcript level of the amiRPDS1_TAVmut- and amiRPDS6-treated tissue remained at the similar level as the pCB301 (empty vector; EV) negative control (Figure 7a). *T*-test was used to determine for statistical significance. Because the TAV vector is unable to cause systemic plant infection without the viral B component, we examined the infiltrated tissues and observed different degrees of mild bleaching that corresponded to the results of NbPDS transcript accumulation levels (Figure 7b).



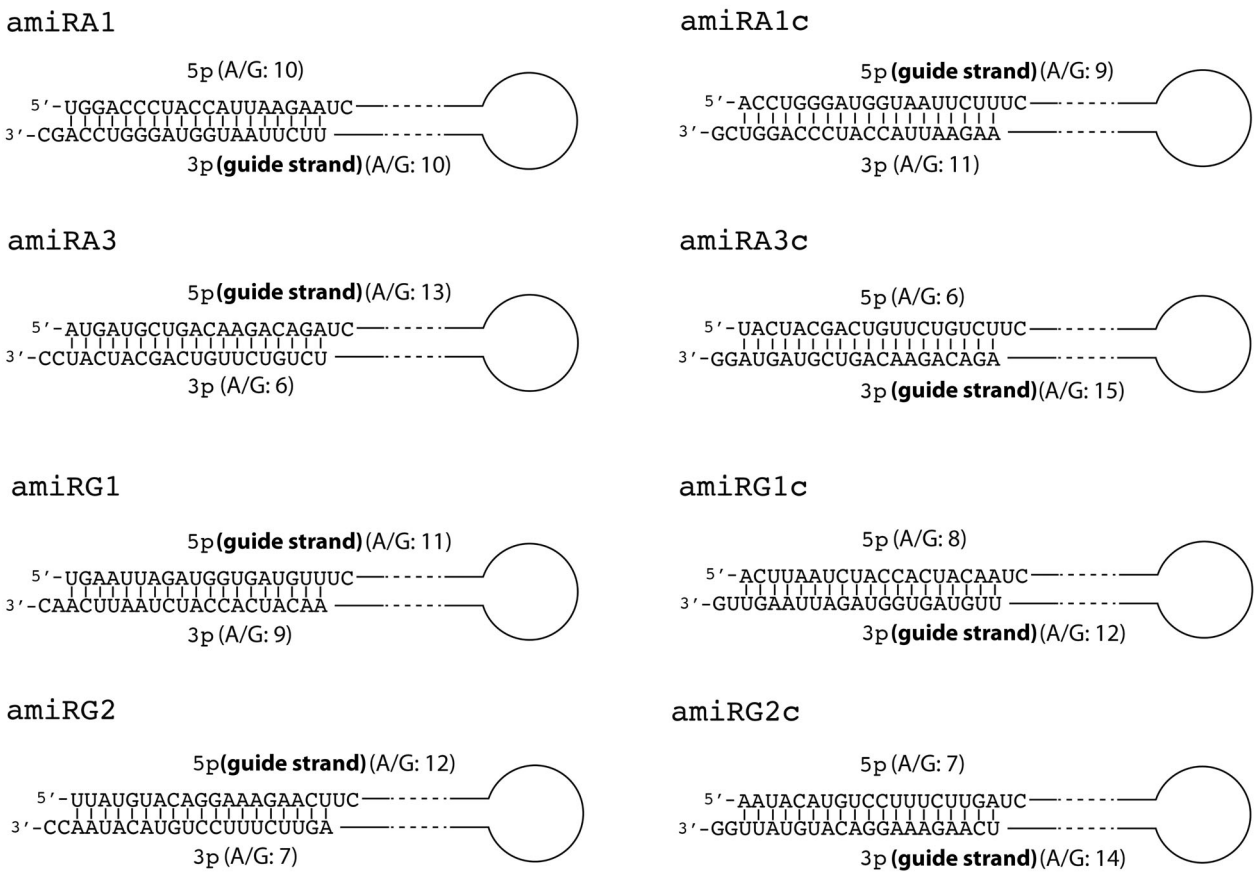


Figure 3 Small RNA northern blot analyses for detecting 5p or 3p strand of TAV-amiRA1, A1c, A3, A3c, G1, G1c, G2, and G2c. Blots were hybridized using probes for a given sequence and exposed to X-ray film. Then blots were stripped and re-probed for the other strand. All probes were of the same specific radioactivity, and the same exposure time were applied to the blot that probed for opposite strands. Higher intensity signals indicate more of the specific amiRNA and are interpreted to be the guide strand. The predicted folding structures show that the sequences in the mature amiRNA duplex contain no mismatches in all the amiRNAs. The strands show higher accumulation levels are enclosed in boxes and considered as the guide strand.

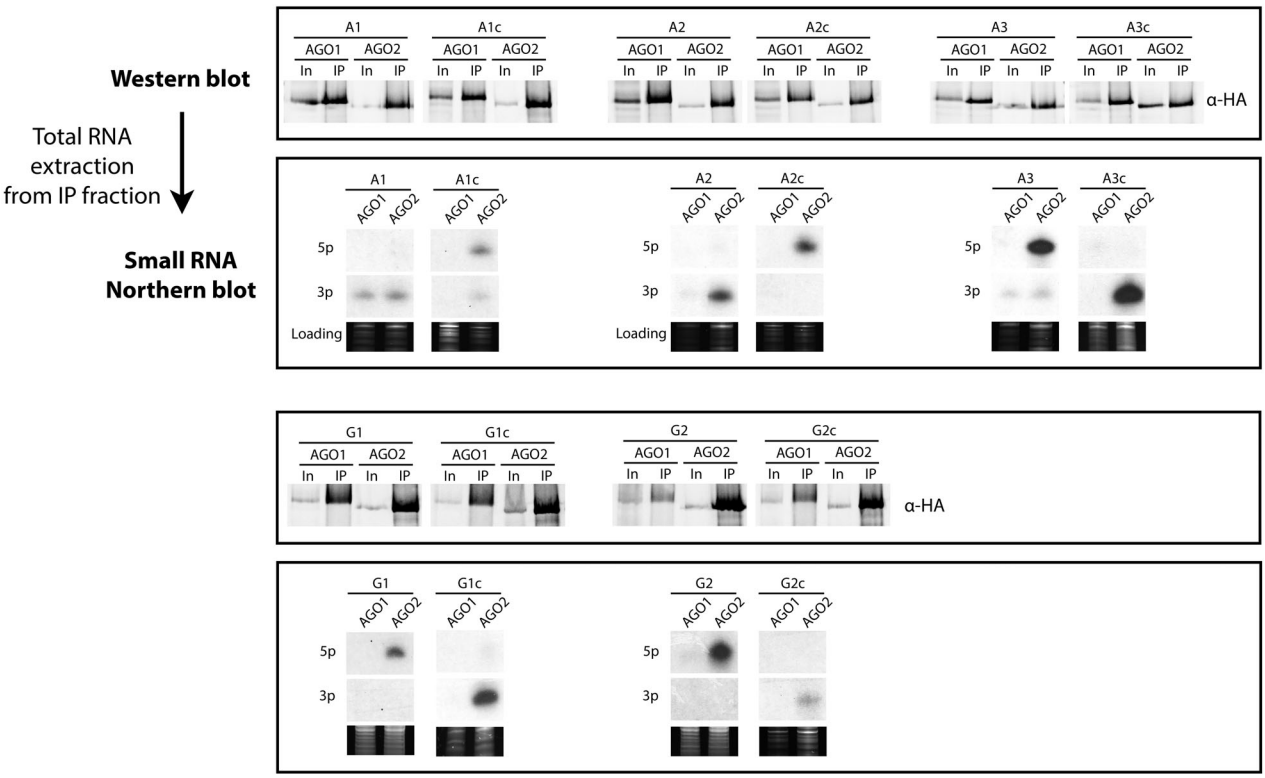
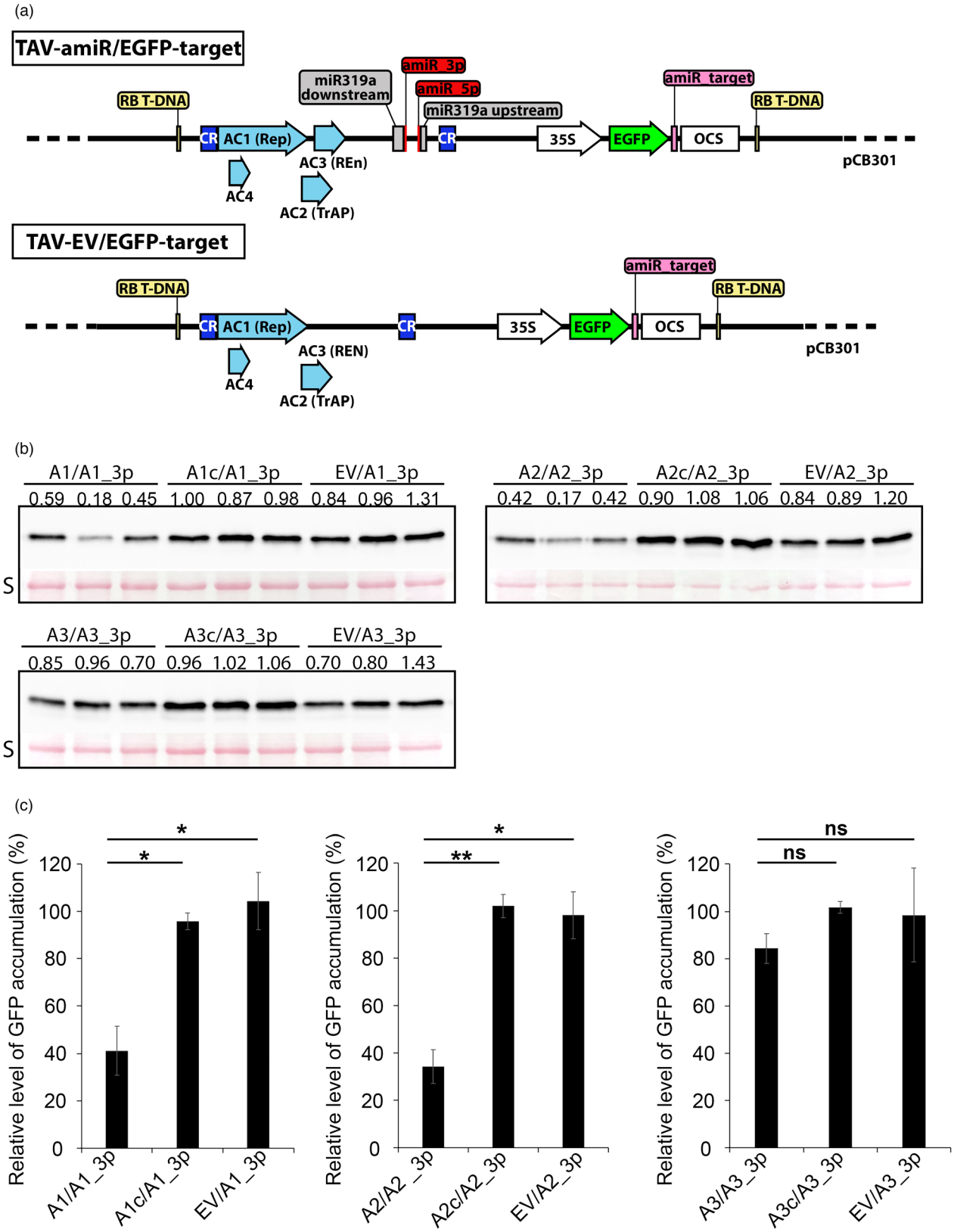


Figure 4 Co-immunoprecipitation results showed that the TAV-amiRNAs predominantly loaded into AGO2 but not AGO1. Western blot analyses showed that both AGO1 and AGO2 proteins were pulled down by HA antibody conjugated beads. The eluted proteins were detected by anti-HA antibodies in the Western blot analyses. 3xHA-AGO1 protein migrated at ~130 kDa and 1xHA-AGO1 migrated at ~120 kDa in size compared with the protein markers used in Western blot analyses. Small RNA northern blot analyses for the RNA extracted from co-IP analyses showed the hybridized signals at the 21-nt position compared with the microRNA ladder. Blots were first probed for 5p strand and, after exposing to X-ray films, blots were stripped and re-probed for the 3p strand. In: Input, IP: immunoprecipitated fraction. A1: TAV-amiRA1, A1c: TAV-amiRA1c, A2: TAV-amiRA2, A2c: TAV-amiRA2c, A3: TAV-amiRA3, A3c: TAV-amiRA3c, G1: TAV-amiRG1, G1c: TAV-amiRG1c, G2: TAV-amiRG2, G2c: TAV-amiRG2c.

Discussion

Using viruses as vehicles for expression of specific amiRNAs in plants offers many advantages over traditional transgenic plant approaches. Many published amiRNA studies showed that the designed amiRNAs, expressed either from non-viral plasmid vector or from an engineered virus, gave the desired effects (Ali *et al.*, 2013; Basso *et al.*, 2019; Niu *et al.*, 2006; Petchthai *et al.*, 2018; Tang *et al.*, 2010). However, in addition to desired phenotype, it is very important also to verify that the intended small RNAs were actually produced from the miRNA pathway. In some other cases, the designed amiRNAs would not function as expected (Arroyo *et al.*, 2014). Therefore, although there are approaches established for screening efficacies of predicted amiRNAs (Li *et al.*, 2013), it is very important to understand why some amiRNAs did not show expected effects and some did and how to ensure that the intended amiRNA strand is selected as a guide RNA to give intended effects. This will lead to more comprehensive understanding of miRNA mechanisms and further improve applications using amiRNA biotechnological applications.

Figure 5 EGFP reporter assays indicated only the correct target sequences corresponding to the amiRNA guide strands triggered silencing effects of the EGFP protein accumulation level. (a) The sequence organization of TAV-amiRNA-EGFP_target and the TAV-EGFP_target (empty vector) constructs. The binary vector pCB301 was used for all the TAV viral vector clones. The sequences of TAV-amiRNA and/or the 35S promotor-driven EGFP with OCS terminator sequences were cloned in between the T-DNA right boarder (RB) and left boarder (LB) sequences of the binary vector. The viral common region (CR) was indicated in blue boxes. The viral genes AC1 (Rep; replication associated protein), AC2 (TrAP; transcriptional activator protein), AC3 (REn; replication enhancer protein) and AC4 were indicated in light blue boxes. The up-stream and down-stream sequences of the pri-miR319a backbone were indicated with grey boxes. The amiRNA-target sequences were cloned at the 3′-UTR of the EGFP sequence. (b) Results of the Western blot analyses. The samples infiltrated with TAV-amiRA1-EGFP_A1_3p (A1/A1_3p) showed silencing effects compared with the accumulation level of the controls: TAV-amiRA1-EGFP_A1c_3p (A1/A1c_3p, non-target) and the TAV-EGFP_A1_3p (empty vector). The same effects were shown in the TAV-amiRA2-EGFP_A2_3p (A2/A2_3p) compared to its controls: TAV-amiRA2-EGFP_A2c_3p (A2/A2c_3p, non-target) and the TAV-EGFP_A2_3p (empty vector). However, contrary to the amiRA1 and amiRA2, both of which the 3p strand was the purine-rich guide strand, and the guide strand of the amiRA3 was the 5p. Therefore, the TAV-amiRA3-EGFP_A3_3p did not show silencing effects compared to the controls: TAV-amiRA3-EGFP_A3c_3p and TAV-EGFP_A3_3p. The bands from the ponceau S staining (S) were used as loading controls. (c) Quantitative analyses of the Western blot analyses results were presented in bar graphs. Statistical significance was analyzed using T-test. **P*-value <0.05; ***P*-value <0.01; ns: non-significant.



Different animal viruses such as *Adenovirus* (dsDNA viral genome), adeno-associated virus (ssDNA viral genome) and bornavirus (borna disease virus (BDV), negative sense (-)ssRNA viral genome) have been developed to express miRNAs in animal systems (Herrera-Carrillo *et al*., 2017; Honda *et al*., 2016). While these diverse viral vectors have their own properties that are more suitable for specific cell types or therapeutic purposes, the replication of all of these includes certain stage(s) within the

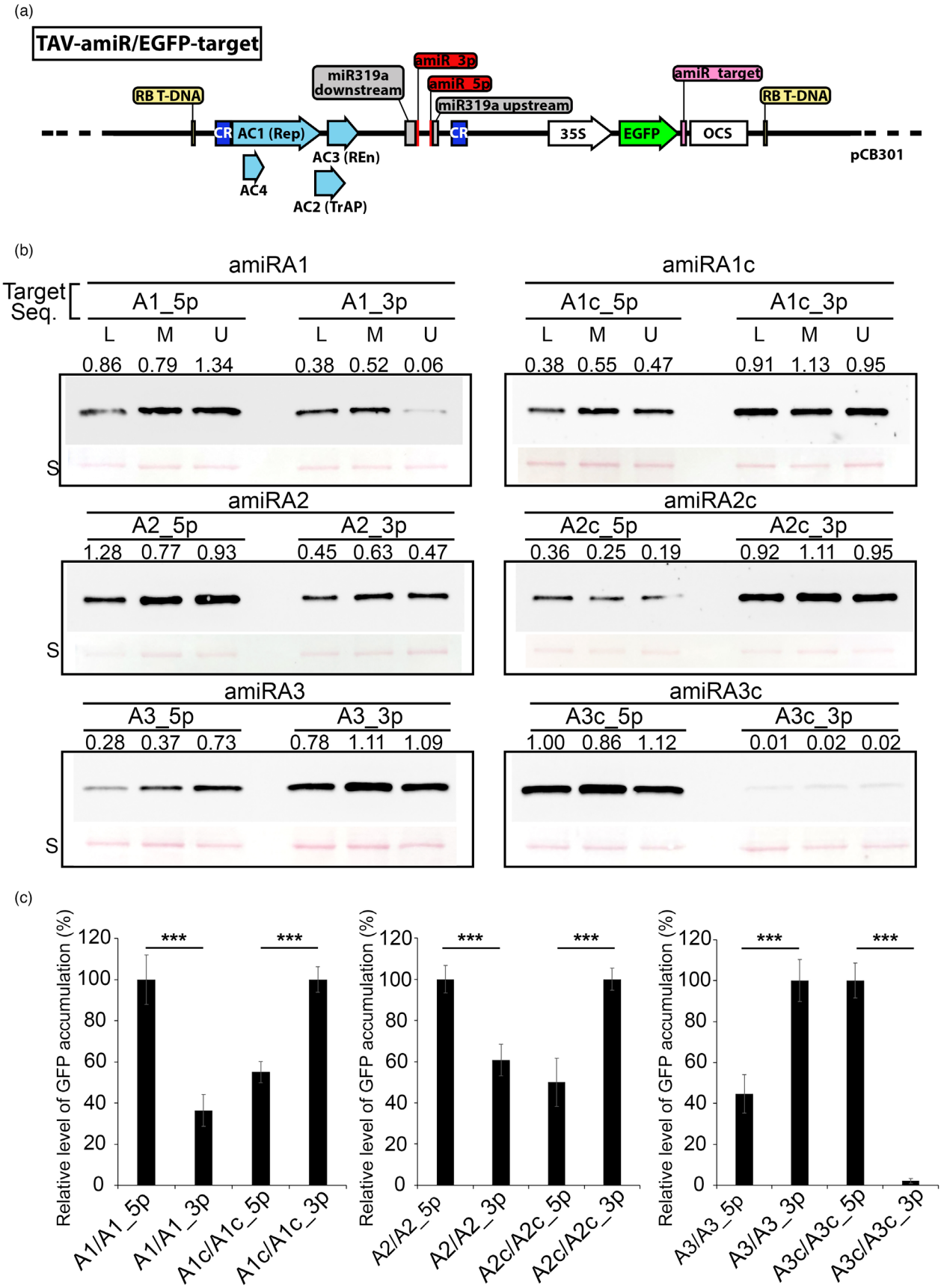

(a)
TAV-amiR/EGFP-target
RB T-DNA
CR
AC1 (Rep)
AC4
AC2 (TrAP)
AC3 (REn)
miR319a downstream
amiR_3p
amiR_5p
miR319a upstream
CR
35S
EGFP
amiR_target
OCS
RB T-DNA
pCB301
(b)
Target Seq.
amiRA1
A1_5p
A1_3p
L M U L M U
0.86 0.79 1.34 0.38 0.52 0.06
amiRA1c
A1c_5p
A1c_3p
L M U L M U
0.38 0.55 0.47 0.91 1.13 0.95
S
amiRA2
A2_5p
A2_3p
1.28 0.77 0.93 0.45 0.63 0.47
amiRA2c
A2c_5p
A2c_3p
0.36 0.25 0.19 0.92 1.11 0.95
S
amiRA3
A3_5p
A3_3p
0.28 0.37 0.73 0.78 1.11 1.09
amiRA3c
A3c_5p
A3c_3p
1.00 0.86 1.12 0.01 0.02 0.02
S
(c)
Relative level of GFP accumulation (%)
120
100
80
60
40
20
0
***
A1/A1_5p
A1/A1_3p
A1c/A1c_5p
A1c/A1c_3p
A2/A2_5p
A2/A2_3p
A2c/A2c_5p
A2c/A2c_3p
A3/A3_5p
A3/A3_3p
A3c/A3c_5p
A3c/A3c_3p

Figure 6 EGFP reporter assays showed that only the target sequence corresponding to the AGO loading guide strand induced silencing effects at the protein level. (a) The constructs of TAV-amiRNA-EGFP_target_5p or _3p were used for these assays. Please see the figure legend of Figure 5 for the details of the sequence/construct information. (b) Representative results of the Western blot analyses. TAV-amiRA1-EGFP_A1_3p (A1/A1_3p) showed silencing effects compared to the results of TAV-amiRA1-EGFP_A1_5p (A1/A1_5p), while TAV-amiRA1c-EGFP-A1c_5p (A1c/A1c_5p) showed silencing effects compared to the results of TAV-amiRA1c-EGFP_A1c_3p (A1c/A1c_3p). Similar results were found in the assays for the TAV-amiRA2-EGFP_A2_3p (A2/A2_3p_ or _5p (A2/A2_5p) and the assays for the TAV-amiRA2c-EGFP_A2c_5p (A2c/A2c_5p) or _3p (A2c/A2c_3p). However, contrary to the amiRA1, amiRA1c and amiRA2, amiRA2c, the purine-rich guide strand of amiRA3 was the 5p strand, and the purine-rich guide strand of the amiRA3c was the 3p strand. Therefore, the TAV-amiRA3-EGFP_A3_5p (A3/A3_5p) showed silencing effects compared to the results of TAV-amiRA3-EGFP_A3_3p (A3/A3_3p), and the TAV-amiRA3c-EGFP_A3c_3p (A3c/A3c_3p) showed silencing effects compared to the results of TAV-amiRA3c-EGFP_A3c_5p (A3c/A3c_5p). U: upper leaf; M: middle leaf; L: lower leaf. The bands from the ponceau S staining (S) were used as loading controls. The complete results of these assays were shown in supplemental Figure 4 (Figure S4). (c) Quantitative analyses of the Western blot analyses results were presented in bar graphs. Statistical significance was analyzed using the *T*-test. ***P-value <0.001.

host cell nucleus. By contrast, in plants, both the nuclear-replicating cabbage leaf curl virus (CaLCuV, a begomovirus) and the cytoplasmic-replicating barley stripe mosaic virus (BSMV, a virgavirus, a (+) ssRNA virus) were reported to be able to express functional amiRNAs based on resulting plant phenotype, although in neither case were the resulting small RNAs confirmed to be authentic amiRNAs (Jian *et al*., 2017; Tang *et al*., 2010). Our studies here showed that although both the nuclear-replicating DNA virus (TAV) and the cytoplasmic-replicating RNA virus (TRBO) were expressed first from binary vector delivery of T-DNAs into the nucleus in plants, the cytoplasmic-replicating RNA virus (TMV-based TRBO vector, which like BSMV is also a virgavirus) did not yield authentic amiRNAs. Rather, the results from our sequencing and northern blot analyses suggest siRNA production from TRBO expression triggered by viral replication. There is a report of a cytoplasmic-replicating virus that can be used for (a)miRNA expression. Recombinant Sindbis virus, rSINV, has been used successfully to express amiRNA sequences originating from constructed pri-miRNA; however, this was not in plants, but in an animal system (Shapiro *et al*., 2010). This study revealed that the rSINV infection resulted in unique re-distribution of Drosha in a virus-specific manner (Shapiro *et al*., 2012). Drosha, a nuclear RNase III enzyme, which is necessary for initial processing of pri-miRNAs, was relocated from the nucleus to the cytoplasm following rSINV infection (Shapiro *et al*., 2012). There is no evidence so far that this kind of phenomenon happens in plant virus infections and thus a nuclear-replicating virus such as used here was necessary to ensure that pri-miRNAs are properly processed in the nucleus.

Both 5p and 3p strands can become the guide strand under different conditions and thus lead to miRNAs with different target sequence specificity (Yang *et al*., 2013; Zinovyeva *et al*., 2015). With an *in vitro* reconstituted system, it was shown that although human AGO2 alone is sufficient for strand selection for certain siRNAs and miRNAs, Dicer and its two other dsRNA binding protein partners, TRBP (*trans*-activation response RNA-binding protein) and protein activator of PKR (PACT), are also often essential for reaching maximal levels of strand selection (Noland and Doudna, 2013). However, not as much has been done to understand factors determining miRNA guide strand selection in plants.

In our studies, we also found that the majority of the TAV-amiRNAs tested by us loaded into AGO2 instead of AGO1, which was found to be the primary AGO protein that mature miR319a was loaded into in *Arabidopsis* (Sobkowiak *et al*., 2012). In flies, the structure of the duplexes appears to be a major determinant for sorting of miRNAs (intrinsically asymmetric in the mature duplex) and siRNAs (symmetric) into AGO1 and AGO2, respectively (Czech *et al*., 2009; Tomari *et al*., 2007), and this may be similar in plants. In addition, Zhang *et al* showed that structural features of miRNA duplexes can also affect miRNA processing and accumulation and found that changes in miRNA duplex structure affected miRNA sorting efficiency into *Arabidopsis* AGO1 or AGO2 (Zhang *et al*., 2014). It was also reported that central mismatches in the miRNA duplex are essential for AGO1 and/or AGO2 sorting (Kawamata *et al*., 2009; Zhang *et al*., 2014). Therefore, designing the amiRNAs containing no mismatches (symmetric) in the mature duplex region may cause the majority

Table 3 The amiRNAs used in the *N. benthamiana* phytoene desaturase silencing assays

amiRNA	Strand	Sequence (5′–3′)	Target site (strand)	# of A/G
amiRPDS1	5p	UUCAGUAUAAAACAUUUGACA	ORF (+)	11
	3p	UCAAAUGUUUUAUACUGAAUA	(–)	10
amiRPDS1_TAVmut	5p	UUCAGUAUAAAACAUUUGACA	ORF (+)	11
	3p	UCAAAUGUUUUAUACUGAAUA	(–)	10
amiRPDS3	5p	ACAUGGCAAUGAACACCUCAU	ORF (+)	11
	3p	GAGGUGUUCAUUGCCAUGUCA	(–)	10
amiRPDS4	5p	AAAUUUUGUGUACAGAAUUAA	3′-UTR (+)	12
	3p	AAUUCTGUACACAAAAUUUAA	(–)	11
amiRPDS5	5p	AGGAGGGUUACCAUCUAAAAA	ORF (+)	14
	3p	UUUAGAUGGUAACCCUCCUGA	(–)	9
amiRPDS6	5p	UUCAUCUUAAAUUUUGUGUAC	3′-UTR (+)	7
	3p	ACACAAAAUUUAAGAUGAAGG	(–)	15

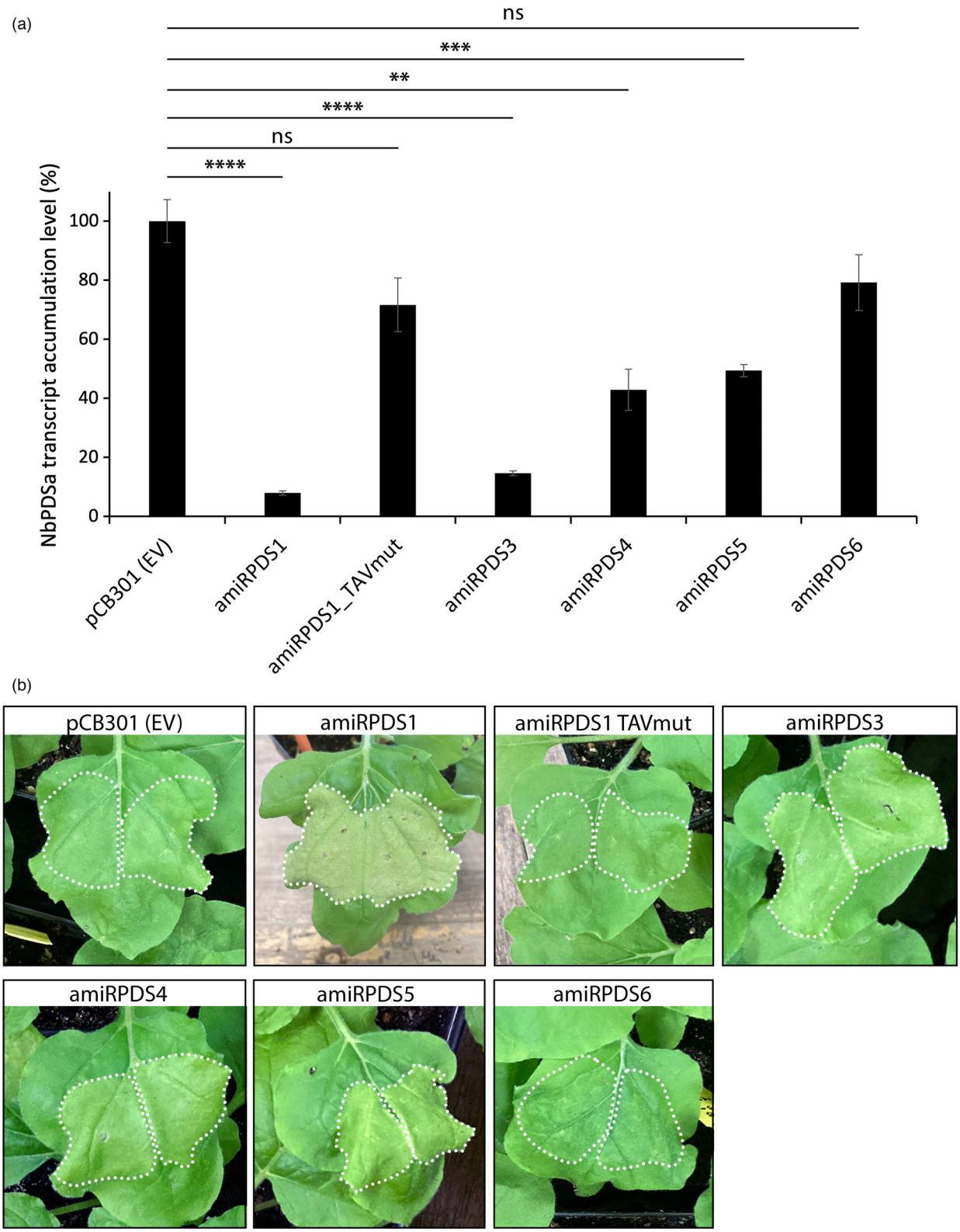


Figure 7 amiRNA strand selection affected silencing effects of the plant endogenous gene, phytoene desaturase (PDS) in *Nicotiana benthamiana*. (a) The results of the RT-qPCR analyses showed decreased RNA transcript accumulation level in the amiRPDS1-, amiRPDS3-, amiRPDS4- and amiRPDS5-treated plants, while the controls, pCB301 (empty vector), amiRPDS1_TAVmut (non-replicable viral vector) and amiRPDS6, of which the purine-rich strand targets the negative strand of the NbPDS gene, showed no (pCB301) or very mild (amiRPDS1_TAVmut and amiRPDS6) silencing effects in the infiltrated tissue. (b) The infiltrated *N. benthamiana* tissue showed different degrees of mild bleaching that corresponded to the results of NbPDS transcript accumulation level compared to the empty vector control.

of amiRNAs to be redirected and loaded into the AGO2. In our experiments, the only exception was the amiRA1, which was loaded into both AGO1 and AGO2 with similar intensity (Figure 4). Moreover, the deep sequencing (next generation sequencing, NGS) analyses of the resulting specific amiRNA guide strands produced from the TAV and pGWB2 showed that the same sequence/strand was processed and selected for both, and the guide sequences could derive from 5p or 3p strand depending on the amiRNA sequence compositions with both vectors (Table 1 and Figure 2e). This suggested that the AGO loading/guide strand selection was not affected by the vectors used to express the up-stream pri-miRNA, and the RNAi effects of the same selected amiRNA sequence were expected to be the same. However, as we show, the different vectors yield different quantities of amiRNAs, which likely affect efficacy. Furthermore, our results of reporter assays also confirmed that only those purine-rich strands that were selected and loaded into AGO proteins induced intended RNAi silencing effects.

Strand selection has been overlooked for the applications using amiRNAs in plants. Our studies not only showed that sequence composition could affect miRNA guide strand selection, but also provide important and useful information for designing mature amiRNAs. The guide strand has been assumed to be the same strand as the guide strand of the endogenous precursor that is used to express the amiRNAs. Our studies revealed that the guide strand of amiRNAs can alter between 5p and 3p arms when sequences of different purine/pyrimidine composition are used even with the same precursor and conditions. Our results indicated that removing the mismatches in the mature amiRNA duplex and designing the intended guide strand to contain excess purines could provide better control of the guide strand selection of amiRNAs for functional RNAi effects. The information provided through our studies will help to improve the application approaches for amiRNAs for use in plants.

Experimental procedures

Construction of amiRNAs

The pri-amiRNA backbone construction was based on the work by Liang *et al*. (2012). The pri-miR319a backbone was amplified from the plasmid pRS300, purchased from addgene (Plasmid #22846) and used for the further cloning. The up-stream and down-stream regions of miR319a were amplified by PCR using the primers MIR319a-up-F-NotI (5′-AAAGCGGCCGCCAAACACACGCTCGGACGCAT-3′), MIR319a-up-R-EcoRI (5′-AAAGAATTCATATATTCCTAAAACATCAATTC-3′), MIR319a-down-F-HindIII (5′-AAAAAGCTTTTTGTATTCCAATTTTCTTGATTAA-3′) and MIR319a-down-R-KpnI (5′-TTTGGTACCCATGGCGATGCCTTAAATAAAG-3′). The amplified products were then cloned into pBluescript SK using the corresponding enzyme sites. To insert the stem-loop region in between the up- and down-stream of the pri-miR319a backbone, the stem-loop region of amiRA2, amiRA2c and amiRG1 were amplified by PCR using the primers amiRATPase319a-EcoRI-2F (5′-AAAGAATTCGTTCTCTGCAGCCTCTATTAATCACAGGTCGTGATATGATTCA-3′), amiRATPase319a-HindIII-2R (5′-AAAAAGCTTGTTCTCTGCAGCCTCTATTAATCAAAGAGAATCAATGATCCA-3′), amiRATPase319a-EcoRI-2cF (5′-AAAGAATTCCAAGAGACGTCGGAGATAATTTCACAGGTCGTGATATGATTCA-3′), amiRATPase319a-HindIII-2cR (5′-AAAAAGCTTCAAGAGACGTCGGAGATAATTTCAAAGAGAATCAATGATCCA-3′), amiRGFP319a-EcoRI-1F (5′-AAAGAATTCGTTGAATTAGATGGTGATGTTTCACAGGTCGTGATATGATTCA-3′) and amiRGFP319a-HindIII-1R (5′-AAAAAGCTTGTTGAATTAGATGGTGATGTTTCAAAGAGAATCAATGATCCA-3′). The underlined sequences denote the mature amiRNA sequences. The complete amiRA2 and amiRA2c backbones were then cloned into the nuclear-replicating viral vector (TAV), the cytoplasmic-replicating viral vector (TRBO) and the non-viral vector pGWB2 with the InFusion Cloning kit (Takara Bio, U.S.A.), following the instructions from the manufacturer. The DNA component A of ToMoV (TAV: ToMoV DNA A Vector), which encodes all the proteins required for viral replication, in the binary vector pCB301 was used for amiRNA expression through agroinfiltration without the DNA B component. In the absence of the B component, which encodes for the viral proteins for cell-to-cell and systemic movement in plant hosts, the A component can only replicate in the infiltrated cells, hence no systemic infection. The cytoplasmic-replicating RNA viral vector-TRBO, a TMV-based viral vector (Lindbo, 2007) was used as well. Similar to the TAV vector, TRBO can only replicate in the infiltrated cells without causing systemic infections.

Agroinfiltration

The C58-C1 strain of *A. tumefaciens* was cultured in 2 mL of LB medium supplemented with kanamycin and grown at 28 °C for 48 h with shaking (250 rpm). Following incubation, a 1-mL aliquot of the culture was transferred to 50 mL of LB with 10-mM MES (pH 5.6) and 20 μL of 100-mM acetosyringone was added. Bacteria were grown at 28 °C for 16–18 h with shaking (250 rpm). After cultures had reached an OD600 = 1, the bacterial suspension was centrifuged at 2520 *g* for 10 min and the pellet was resuspended in 50-mL $MgCl_2$ buffer (10 mM $MgCl_2$). Seventy-five microlitres of 100 mM acetosyringone was added, and the bacterial suspension was kept at room temperature for at least 3 h without shaking before infiltration. P19, the silencing suppressor of tomato bushy stunt virus, alone was used as a negative control in the small RNA northern blot and small RNA deep sequencing analyses in Figure 2.

RNA extraction and small RNA northern blot analysis

Infiltrated tissues were collected 3 dpi, and RNA was extracted with TRIzol (Invitrogen) according to the manufacturer's instructions, except that the RNA was not washed with 70% ethanol before resuspending in 120-μL RNase-free water. The resuspended total RNA was treated with 40 ml of 5 M NaCl and 40 μL of 50% PEG 8000 to each sample to enrich for small RNAs. After 30 min on ice, the mixture was centrifuged at 12,000 *g* at 4 °C for 10 min. The pellet that contained the large RNA was washed with 75% EtOH and resuspended in 30 μL of RNase-free water (the large RNA was used for determining viral replication of TRBO). The supernatant was then transferred to a new tube and 3X volume of cold 100% EtOH was added. The solution was incubated in −20 °C overnight and then centrifuged 12,000 *g* at 4°C for 20 min. The pellet was dried and resuspended in 30-μL RNase-free water. The RNA concentration was determined with a NanoDrop spectrophotometer (ND-1000), and 20 μg of RNA/sample was fractionated in 7-M urea denaturing 15% polyacrylamide gels (15% acrylamide, 1X TBE, 7 M urea). microRNA ladder (NEB) was used for the small RNA northern blot analyses as a standard for small RNA sizes. Equal loading and integrity of the RNA was determined by visualization of ethidium bromide-stained gels. The RNA was transferred to an Amersham Hybond-N^+ membrane (Amersham) with a semi-dry transfer apparatus (Bio-Rad). Membranes were prehybridized for 3 h in ULTRAhyb-Oligo hybridization buffer (Ambion), followed by hybridization in

the same buffer with a dAT^{32}P-labelled probes at 42 °C for 16 h. Probe labelling was done by 5′-end labelling with T4 polynucleotide kinase (NEB). Membranes were washed twice with 2X SSC solution containing 0.2% SDS at 42 °C for 30 min. Hybridization signals were visualized by autoradiography.

In the analysis for purine-rich preference in guide strand selection, the blots were stripped after the first hybridization for detecting 5p strand with 2% SDS in RNase-free water and incubated in 90-100 °C for 1 h. After stripping, the blots were re-prehybridized and followed by the hybridization steps described above to detect the complementary strand.

Northern blot analysis for viral replication of TRBO clones

Two micrograms of the large RNA fraction was used for confirming viral replication of TRBO. 1% HEPES-EDTA agarose gels were stained with SYBR Gold Nucleic Acid Gel Stain (Thermo Fisher Scietific) to show the RNA loading. The RNA was transferred to an Amersham Hybond-N^{+} membrane (Amersham) by capillary blotting. Membranes were prehybridized for 3 h in NorthernMax™ Prehyb/Hyb buffer (Invitrogen), followed by hybridization in the same buffer with a dUTP-^{32}P-labled probes at 65 °C for 16 h. Probe labelling was done by *in vitro* transcription with the MAXIscript™ T7 Transcription Kit (Ambion). Membranes were washed twice with 2X SSC solution containing 0.2% SDS at 42 °C for 15 min each time and followed by twice with 2X SSC solution containing 1% SDS and 2% SDS, respectively, at 68 °C for 30 min each time. Hybridization signals were visualized by autoradiography.

PCR analysis for viral replication of TAV clones

The agroinfiltrated leaf tissues were collected 3 days post infiltration (dpi). Total DNA was extracted with Dneasy Plant Mini Kit (Qiagen) following the instructions from the manufacturer. The primers used for confirming recombinant ToMoV replication were TAV AC3 sq F (5′-ATGTAAATCATGTAT TGGAGAATCATTCAATAAAATTCA-3′) and TAV PW 8 (5′-TGATTGCCAATCTTTCTGGG-3′). The product size indicating replication was 1853 base pairs (bp). If not replicating, the PCR product would contain sequence of the binary vector pCB301 with the size of 6172-bp.

Illumina small RNA library preparation

The small RNA libraries were prepared following the instructions from the manufacturer. In short, 1 μg of the small RNA-enriched fractions, isolated with mirVana miRNA isolation kit (Ambion/Thermo Fisher Scientific), was used for the library preparation. The RNAs were then used for the sequential ligation of the RNA 3′ and RNA 5′ RNA adapters (indexes). The ligations were followed by reverse transcription (RT)-polymerase chain reaction (PCR) to create cDNA constructs based on the small RNA ligated with 3′ and 5′ adapters. PCR was performed with two primers that annealed to the ends of the adapters. The cDNAs were gel (5% Mini-PROTEAN TBE precast gel, Bio-Rad) purified, eluted and ethanol precipitated. The cDNAs were resuspended in 10-μL 10-mM Tris-HCl, pH 8.5. The cDNA libraries were validated by Experion Automated Electrophoresis system (Bio-Rad). The validated cDNA libraries were sent to QB3 Vincent J. Coates Genomics Laboratory at the University of California, Berkeley for HiSeq Illumina deep sequencing (HiSeq 2000 platform). This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303.

Bioinformatic analysis

The deep sequencing reads were trimmed with CLC Genomics Workbench. The graphics were produced by using the R program with viRome package. Specific sequence reads were extracted and counted by python and analyzed with total reads with length of 18- to 24-nt. Multiple pairwise tests against a reference group and multiple grouping variables were used to analyze reporter assays in R. The ANOVA and *t*-test were used for determining statistical significance levels.

Construction of HA-tagged AGO1 and AGO2 expression plasmids

The 3x HA tag DNA sequence was cloned at the 5′-end of the AGO1 open reading frame (ORF) with 1 threonine and 1 serine between each HA and between HA and AGO1 protein as linkers. The 3xHA-AGO1 was cloned into the expression vector pdM, which contains the 2x35S promoter and the OCS terminator. The sequence of the 2x35S promoter, 3xHA-AGO1 and the OCS terminator was then cloned into the binary vector pART27 with the Not I restriction enzyme sites for *Agrobacterium* infiltration. The 1x HA AGO2 was cloned into the binary vector with the same strategy as the 3xHA-AGO1 construct. Both 3xHA-AGO1 and 1xHA-AGO2 constructs were transformed into the GV3101 *Agrobacterium tumefaciens* strain.

Co-immunoprecipitation (co-IP)

The protocol for co-IP was adopted from the protocol published by Zhao *et al*. (2012). ~4 g of infiltrated *N. benthamiana* leaf tissue for each sample was collected and was ground with liquid nitrogen. The ground tissue was transferred to a centrifuge tube with 4 mL of extraction buffer [20-mM Tris-HCl, 300-mM NaCl, 5-mM $MgCl_2$, 0.5% (v/v) NP40, and 1 tablet of Pierce Protease Inhibitor and 200 μL of Halt Protease Inhibitor Cocktail, EDTA-free (Thermo Fisher Scientific) in each 50 mL volume]. The ground samples in the extraction buffer were thawed on ice and shaken for 15 min. The samples were then centrifuged at 10 000 *g*, 4 °C for 10 min. The supernatant was filtered through a 100-μm cell strainer. Sixty microlitres of filtrate was collected as input samples. Sixty microlitres of Pierce™ anti-HA magnetic beads (Thermo Fisher Scientific) was washed with extraction buffer and added into the rest of the filtrate. The mixture was gently rotated at 10 °C for 2 h. After the incubation, the beads were washed five times with the washing buffer (20-mM Tris-HCl, 300-mM NaCl, 5-mM $MgCl_2$, 0.5% (v/v) triton X-100, and the same protease inhibitors as used in the extraction buffer). 1/5 of the beads were collected for Western blot analyses. The rest of the beads were used for total RNA extraction.

RNA extraction and northern blot analysis for co-immunoprecipitation samples

The RNA extraction for co-IP samples was done following the protocol published by Zhao *et al*. (2012).

The concentration of the purified RNA was measured by Qubit 4 fluorometer (Thermo Fisher Scientific). Equal amount of each RNA sample was divided for 2 northern blot analyses and probed to test for the 5p or 3p arm of each amiRNA. Regular or locked nucleic acid (LNA) modified single stranded DNA oligos complementary sequence for each strand of each amiRNA were synthesized. The LNA modified probes were used for low signal amiRNAs from the co-IP assays (amiRA3c, amiRG1 and amiRG2). 4×10^6 μCi probes was added per mL of hybridization buffer.

The RNA was hybridized for at least 16 h and washed by 2 × SSC solution containing 0.1% SDS twice at 42 °C for 30 min. Hybridization signals were visualized by autoradiography.

Reporter assays

The 35S promoter-driven EGFP ORF with 54 nt amiRNA-target sequences containing the 21-nt expected target sequences linked at the 3′ UTR was cloned in the TAV-amiRNA clones accordingly. Target sequences of 5p or 3p strand of each amiRNA were cloned to evaluate the RNAi effects of both strands of each tested amiRNA. The clones were agroinfiltrated into *N. benthamiana* plants, which were kept in a 24 °C, 16-h light-conditioned growth chamber. Three samples were collected from different leaves of each tested plant and three plants were tested (total nine samples) for each amiRNA-target combination at 4 dpi.

Western blot analysis

One-fifth of the beads collected from co-IP assays were mixed with 60 μL of sample loading buffer (2% SDS, 1% beta-mercaptoethanol, 10% sucrose, 0.005% bromophenol blue, 0.1M Tris-HCl pH 6.8) and boiled for 10 min to release the bound proteins from the beads. After removing the beads, the samples were loaded into 10% acrylamide gels along with the PageRuler Prestained Protein ladder (Thermo Fisher Scientific). The proteins were then transferred to nitrocellulose membranes. Membrane blots were incubated with anti-HA.11 epitope tag antibodies (BioLegend) diluted at 1:2000 dilution in the TSW buffer (10 mM, 0.9% NaCl, 0.25% gelatin, 0.1% Triton X-100 and 0.02% SDS) and with the secondary antibody (HRP-conjugated anti–mouse [1:3000 dilution]) (Bio-Rad). The secondary antibody was detected with the SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific) and visualized with ChemiDoc imaging system (Bio-Rad).

PDS silencing assays

The amiRPDS clones were designed based on the *N. benthamiana* PDS mRNA sequence published on NCBI (Accession number EU165355.1). The mature amiRNA sequences of the amiRPDS constructs were listed in Table 3. Total RNA was extracted from the samples collected at 4 dpi using Trizol reagent (Invitrogen). The extracted RNA was treated with DNase treatment (Promega). Reverse transcription (RT) of the total RNA was done using iScript Reverse Transcription Supermix for RT-qPCR (Bio-Rad). The cDNA was then used for qPCR assays using iQ™ Multiplex Powermix (Bio-Rad). The reactions were done in CFX96™ Real-Time System (Bio-Rad). No-RT controls were included in the qPCR assays. Relative and absolute quantification were used for the accumulation quantification, and both showed the same efficacy results. The primers and probes of NbPDS and the reference COX mRNAs used for RT-qPCR were listed in Table S1. The primers and probe for COX mRNA were designed based on a previous publication (Hughes *et al*., 2006). The plasmid TOPO-Zero-Blunt-NbPDS was used for absolute quantification. The results shown in Figure 5 was the results of relative quantification using COX mRNA as a reference.

Acknowledgements

This work was supported by the Citrus Research Board in California (CRB project No. 5500-198 and 5300-169) and the United States Department of Agriculture-National Institute of Food and Agriculture (USDA-NIFA, Grant No. 13166985). We thank Dr. Jared Nigg and Anneliek ter Horst for providing valuable knowledge and techniques for bioinformatic analysis. We also thank Dr. Robert L. Gilbertson for kindly providing us the TAV vector for our studies.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Y-W Kuo designed and carried out the experiments; BW Falk provided advice and guidance for the project; Y-W Kuo and BW Falk wrote the manuscript.

References

Ali, I., Amin, I., Briddon, R.W. and Mansoor, S. (2013) Artificial microRNA-mediated resistance against the monopartite begomovirus Cotton leaf curl Burewala virus. *Virol J*. **10**, 231.

Arroyo, J.D., Gallichotte, E.N. and Tewari, M. (2014) Systematic design and functional analysis of artificial microRNAs. *Nucleic Acids Res*, **42**, 6064–6077.

Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.

Basso, M.F., Ferreira, P.C.G., Kobayashi, A.K., Harmon, F.G., Nepomuceno, A.L., Molinari, H.B.C. and Grossi-de-Sa, M.F. (2019) MicroRNAs and new biotechnological tools for its modulation and improving stress tolerance in plants. *Plant Biotechnol J*, **17**, 1482–1500.

Carbonell, A., Carrington, J.C. and Daros, J.A. (2016) Fast-forward generation of effective artificial small RNAs for enhanced antiviral defense in plants. *RNA Dis*, **3**.

Carbonell, A., Lopez, C. and Daros, J.A. (2019) Fast-Forward Identification of Highly Effective Artificial Small RNAs Against Different Tomato spotted wilt virus Isolates. *Mol Plant Microbe Interact*, **32**, 142–156.

Czech, B., Zhou, R., Erlich, Y., Brennecke, J., Binari, R., Villalta, C., Gordon, A. *et al*. (2009) Hierarchical rules for Argonaute loading in Drosophila. *Mol Cell*, **36**, 445–456.

Fahlgren, N., Hill, S.T., Carrington, J.C. and Carbonell, A. (2016) P-SAMS: a web site for plant artificial microRNA and synthetic trans-acting small interfering RNA design. *Bioinformatics*, **32**, 157–158.

Fellmann, C., Zuber, J., McJunkin, K., Chang, K., Malone, C.D., Dickins, R.A., Xu, Q. *et al*. (2011) Functional identification of optimized RNAi triggers using a massively parallel sensor assay. *Mol Cell*, **41**, 733–746.

Griffiths-Jones, S., Hui, J.H., Marco, A. and Ronshaugen, M. (2011) MicroRNA evolution by arm switching. *EMBO Rep*, **12**, 172–177.

Herrera-Carrillo, E., Liu, Y.P. and Berkhout, B. (2017) Improving miRNA Delivery by Optimizing miRNA Expression Cassettes in Diverse Virus Vectors. *Hum Gene Ther Methods*, **28**, 177–190.

Honda, T., Yamamoto, Y., Daito, T., Matsumoto, Y., Makino, A. and Tomonaga, K. (2016) Long-term expression of miRNA for RNA interference using a novel vector system based on a negative-strand RNA virus. *Sci Rep*, **6**, 26154.

Hou, Y.M. and Gilbertson, R.L. (1996) Increased pathogenicity in a pseudorecombinant bipartite geminivirus correlates with intermolecular recombination. *J Virol*, **70**, 5430–5436.

Hu, H.Y., Yan, Z., Xu, Y., Hu, H., Menzel, C., Zhou, Y.H., Chen, W. *et al*. (2009) Sequence features associated with microRNA strand selection in humans and flies. *BMC Genom*., **10**, 413.

Hughes, K.J., Tomlinson, J.A., Griffin, R.L., Boonham, N., Inman, A.J. and Lane, C.R. (2006) Development of a One-Step Real-Time Polymerase Chain Reaction Assay for Diagnosis of Phytophthora ramorum. *Phytopathology*, **96**, 975–981.

Hutcheson, J.M., Susta, L., Stice, S.L., Afonso, C.L. and West, F.D. (2015) Delayed Newcastle disease virus replication using RNA interference to target the nucleoprotein. *Biologicals*, **43**, 274–280.

Jian, C., Han, R., Chi, Q., Wang, S., Ma, M., Liu, X. and Zhao, H. (2017) Virus-based MicroRNA silencing and overexpressing in common wheat (*Triticum aestivum* L.). *Front. Plant Sci*. **8**, 500.

Kawamata, T., Seitz, H. and Tomari, Y. (2009) Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat. Struct. Mol. Biol*. **16**, 953–960.

Li, J.F., Chung, H.S., Niu, Y., Bush, J., McCormack, M. and Sheen, J. (2013) Comprehensive protein-based artificial microRNA screens for effective gene silencing in plants. *Plant Cell*, **25**, 1507–1522.

Li, S.C., Liao, Y.L., Ho, M.R., Tsai, K.W., Lai, C.H. and Lin, W.C. (2012a) miRNA arm selection and isomiR distribution in gastric cancer. *BMC Genom*. **13** (Suppl 1), S13.

Li, S.C., Tsai, K.W., Pan, H.W., Jeng, Y.M., Ho, M.R. and Li, W.H. (2012b) MicroRNA 3′ end nucleotide modification patterns and arm selection preference in liver tissues. *BMC Syst Biol*. **6**(Suppl 2), S14.

Liang, G., He, H., Li, Y. and Yu, D. (2012) A new strategy for construction of artificial miRNA vectors in Arabidopsis. *Planta*, **235**, 1421–1429.

Lindbo, J.A. (2007) TRBO: a high-efficiency tobacco mosaic virus RNA-based overexpression vector. *Plant Physiol*. **145**, 1232–1240.

Meijer, H.A., Smith, E.M. and Bushell, M. (2014) Regulation of miRNA strand selection: follow the leader? *Biochem. Soc. Trans*. **42**, 1135–1140.

Motavaf, M., Safari, S. and Alavian, S.M. (2014) Targeting microRNA-122: walking on cutting edge of hepatitis C virus infection therapy. *Acta Virol*. **58**, 301–308.

Niu, Q.W., Lin, S.S., Reyes, J.L., Chen, K.C., Wu, H.W., Yeh, S.D. and Chua, N.H. (2006) Expression of artificial microRNAs in transgenic *Arabidopsis thaliana* confers virus resistance. *Nat. Biotechnol*. **24**, 1420–1428.

Noland, C.L. and Doudna, J.A. (2013) Multiple sensors ensure guide strand selection in human RNAi pathways. *RNA*, **19**, 639–648.

Ohanian, M., Humphreys, D.T., Anderson, E., Preiss, T. and Fatkin, D. (2013) A heterozygous variant in the human cardiac miR-133 gene, MIR133A2, alters miRNA duplex processing and strand abundance. *BMC Genet*. **14**, 18.

Ossowski, S., Schwab, R. and Weigel, D. (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. *Plant J*. **53**, 674–690.

Petchthai, U., Yee, C.S.L. and Wong, S.M. (2018) Resistance to CymMV and ORSV in artificial microRNA transgenic *Nicotiana benthamiana* plants. *Sci Rep*. **8**, 9958.

Reis, R.S., Hart-Smith, G., Eamens, A.L., Wilkins, M.R. and Waterhouse, P.M. (2015) Gene regulation by translational inhibition is determined by Dicer partnering proteins. *Nat. Plants*, **1**, 14027.

Rogers, K. and Chen, X. (2013) Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell*, **25**, 2383–2399.

Rubio, M., Montanez, R., Perez, L., Milan, M. and Belles, X. (2013) Regulation of atrophin by both strands of the mir-8 precursor. *Insect Biochem. Mol. Biol*. **43**, 1009–1014.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199–208.

Shapiro, J.S., Langlois, R.A., Pham, A.M. and Tenoever, B.R. (2012) Evidence for a cytoplasmic microprocessor of pri-miRNAs. *RNA*, **18**, 1338–1346.

Shapiro, J.S., Varble, A., Pham, A.M. and Tenoever, B.R. (2010) Noncanonical cytoplasmic processing of viral microRNAs. *RNA*, **16**, 2068–2074.

Sobkowiak, L., Karlowski, W., Jarmolowski, A. and Szweykowska-Kulinska, Z. (2012) Non-canonical processing of Arabidopsis pri-miR319a/b/c generates additional microRNAs to target one RAP2.12 mRNA isoform. *Front. Plant Sci*. **3**, 46.

Tang, Y., Wang, F., Zhao, J., Xie, K., Hong, Y. and Liu, Y. (2010) Virus-based microRNA expression for gene functional analysis in plants. *Plant Physiol*. **153**, 632–641.

Tomari, Y., Du, T. and Zamore, P.D. (2007) Sorting of *Drosophila* small silencing RNAs. *Cell*, **130**, 299–308.

Voinnet, O. (2009) Origin, biogenesis, and activity of plant microRNAs. *Cell*, **136**, 669–687.

Wagaba, H., Patil, B.L., Mukasa, S., Alicai, T., Fauquet, C.M. and Taylor, N.J. (2016) Artificial microRNA-derived resistance to Cassava brown streak disease. *J. Virol. Methods*, **231**, 38–43.

Warf, M.B., Johnson, W.E. and Bass, B.L. (2011) Improved annotation of *C. elegans* microRNAs by deep sequencing reveals structures associated with processing by Drosha and Dicer. *RNA*, **17**, 563–577.

Xu, L.H., Guo, Y., Zhang, X.L., Chen, J.J. and Hu, S.Y. (2016) Blood-based circulating microRNAs are potential diagnostic biomarkers for leukemia: result from a meta-analysis. *Cell Physiol. Biochem*. **38**, 939–949.

Yang, X., Du, W.W., Li, H., Liu, F., Khorshidi, A., Rutnam, Z.J. and Yang, B.B. (2013) Both mature miR-17-5p and passenger strand miR-17-3p target TIMP3 and induce prostate tumor growth and invasion. *Nucleic Acids Res*. **41**, 9688–9704.

Zhang, S., Liu, Y. and Yu, B. (2015) New insights into pri-miRNA processing and accumulation in plants. *Wiley Interdiscip. Rev. RNA*, **6**, 533–545.

Zhang, X., Niu, D., Carbonell, A., Wang, A., Lee, A., Tun, V., Wang, Z. *et al*. (2014) ARGONAUTE PIWI domain and microRNA duplex structure regulate small RNA sorting in Arabidopsis. *Nat. Commun*. **5**, 5468.

Zhao, H., Lii, Y., Zhu, P. and Jin, H. (2012) Isolation and profiling of protein-associated small RNAs. *Methods Mol. Biol*. **883**, 165–176.

Zhao, Q., Deng, S., Wang, G., Liu, C., Meng, L., Qiao, S., Shen, L. *et al*. (2016) A direct quantification method for measuring plasma MicroRNAs identified potential biomarkers for detecting metastatic breast cancer. *Oncotarget*, **7**, 21865–21874.

Zinovyeva, A.Y., Veksler-Lublinsky, I., Vashisht, A.A., Wohlschlegel, J.A. and Ambros, V.R. (2015) *Caenorhabditis elegans* ALG-1 antimorphic mutations uncover functions for Argonaute in microRNA guide strand selection and passenger strand disposal. *Proc. Natl Acad. Sci. USA*, **112**, E5271–5280.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Small RNA mapping of amiRA2- and amiRA2c-TRBO.
Figure S2 Small RNA mapping of amiRA2- and amiRA2c-pGWB2.
Figure S3 Small RNA mapping of amiRA2- and amiRA2c-TAV.
Figure S4 EGFP reporter assays showed silencing effects of amiRNAs with corresponding target sequences.
Table S1 The primers and probes used in the RT-qPCR analyses for the NbPDS silencing assays.