miR3954 is a trigger of phasiRNAs that affects flowering time in citrus

Yuanlong Liu¹,#, Lili Ke¹,#, Guizhi Wu¹, Yuantao Xu¹, Xiaomeng Wu¹, Rui Xia², Xiuxin Deng¹ and Qiang Xu¹,*

¹: Key Laboratory of Horticultural Plant Biology of Ministry of Education, Huazhong Agricultural University, Wuhan, 430070, China
²: State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Guangzhou, 510642, China

#: These authors contributed equally to this work.

Author for correspondence:

Qiang Xu

Tel: +86-27-87286623
Fax: +86-27-87282010
Email: xuqiang@mail.hzau.edu.cn

Author email addresses:

Yuanlong Liu: yuanlongliu0115@gmail.com

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/tpj.13650
This article is protected by copyright. All rights reserved.
Running title: miR3954 and phasiRNAs affect citrus flowering

Key words: Citrus sinensis, miR3954, phasiRNAs, NAC transcription factor, regulatory network.

Summary

In plant, a few 22 nt miRNAs direct cleavages of their targets and trigger the biogenesis of phased small interfering RNAs (phasiRNAs) in plant. In this study, we characterized a miRNA triggering phasiRNAs generation, miR3954, and explored its downstream target genes and potential function. Our results demonstrated that miR3954 showed specific expression in the flowers of citrus species and it targeted a NAC transcription factor (Cs7g22460) and two non-coding RNA transcripts (lncRNAs, Cs1g09600 and Cs1g09635). The production of phasiRNAs was detected from transcripts targeted by miR3954 and was further verified in both sequencing data and transient expression experiments. PhasiRNAs derived from the two lncRNAs targeted not only miR3954-targeted NAC gene but also
additional NAC homologous genes. No homologous genes of these two lncRNAs was found in plants other than citrus species, implying that this miR3954-lncRNAs-phasiRNAs-NAC pathway is likely citrus-specific. Transgenic analysis indicated that the miR3954-overexpressing lines showed decreased transcripts of lncRNA, elevated abundance of phasiRNAs and reduced expression of NAC genes. Interestingly, the overexpression of miR3954 leads to early flowering in citrus plants. In summary, our results illustrated a model of the regulatory network of miR3954-lncRNA-phasiRNAs-NAC, which may be functionally involved in flowering in citrus.

Introduction

In plants, microRNAs (miRNAs) with a length of 20-24 nucleotides (nt) play vital roles in development. Generally, miRNAs in plants with different lengths are generated through diverse biogenesis pathways and have different functions. In most cases, canonical miRNAs, which are generated by Dicer-like 1 (DCL1), cooperate with ARGONAUTE 1 (AGO1) proteins and constitute RNA-induced silencing complexes (RISCs) to guide posttranscriptional gene silencing through direct cleavage or translation inhibition of target mRNAs (Bartel, 2004, Vaucheret, 2008, Voinnet, 2009, Wang et al., 2011). Phased siRNAs (phasiRNAs), another major types of small RNAs in plants, were firstly reported in the model plant Arabidopsis which were termed as trans-acting siRNAs (tasiRNA) initially, as these siRNAs are derived from non-coding transcripts and function in trans to target downstream genes (Peragine et al., 2004, Vazquez et al., 2004, Allen et al., 2005, Yoshikawa et al., 2005). In Arabidopsis, miR173 triggers the formation of tasiRNAs targeting pentatricopeptide repeat
protein coding genes from \textit{TAS1} and \textit{TAS2}; miR828 targets \textit{TAS4} and triggers the formation of tasiRNAs targeting \textit{MYB} transcription factors (Rajagopalan et al., 2006). miR390 triggers the biogenesis of tasiRNAs from \textit{TAS3}, targeting genes encoding auxin response factor (ARF) (Allen et al., 2005).

Although a few 21 nt miRNAs, such as miR390 and miR161, were loaded into AGO7 and triggered the biogenesis of phasiRNAs (Axtell et al., 2006, Montgomery et al., 2008, Allen and Howell, 2010, Fei et al., 2013), it has been widely reported that 22 nt miRNAs play a dominant role in triggering the biogenesis of phasiRNAs (Allen et al., 2005, Allen and Howell, 2010, Chen et al., 2010, Cuperus et al., 2010). Most 22 nt miRNAs are derived from asymmetric miRNA/miRNA* duplexes that are likely processed by DCL1 (Cuperus et al., 2010). After the direction of accurate slicing, 22 nt miRNAs cooperate with AGO1 to recruit RNA-dependent RNA polymerase 6 (RDR6) to convert the 3’ cleaved fragment into double stranded RNA (dsRNA) (Cuperus et al., 2010). Subsequently, DCL4 processes the dsRNA into 21 nt or 24 nt phasiRNAs in a sequential head-to-tail manner right after the miRNA cleavage site (Allen et al., 2005, Xie et al., 2005, Yoshikawa et al., 2005, Johnson et al., 2009, Allen and Howell, 2010, Cuperus et al., 2010, Yoshikawa et al., 2013, Zhai et al., 2015).

Experimental evidence has demonstrated that increasing the length of a miRNA from 21 nt to 22 nt gave the miRNA a competence of triggering phasiRNAs biogenesis from its target genes (Chen et al., 2010, Cuperus et al., 2010).

With the biogenesis process of phasiRNAs essentially worked out, researchers developed bioinformatics approaches to detect extensive \textit{PHAS} loci of phasiRNAs production and their miRNA triggers in several species, such as soybean, citrus, Norway spruce, and strawberry.
(Xia et al., 2013, Arikit et al., 2014, Liu et al., 2014b, Wu et al., 2015, Xia et al., 2015a, Xia et al., 2015b). To date, studies have indicated that phasiRNAs are involved in different biological processes, such as heat stress (Li et al., 2014), plant defense responses (Zhai et al., 2011, Li et al., 2012, Fei et al., 2013, Zhao et al., 2015), secondary metabolite synthesis (Xia et al., 2012, Rock, 2013), and especially plant development (Peragine et al., 2004, Xie et al., 2005, Liu et al., 2007, Johnson et al., 2009, Cho et al., 2012, Lam et al., 2015, Zhai et al., 2015). In *Physcomitrella patens*, miR156 and miR390 co-target a PHAS transcript and co-modulate the accumulation of phasiRNAs that affect development timing (Cho et al., 2012). After a combined data set of small RNAs from different tissues in rice was analyzed, 831 out of 1029 phasiRNAs were detected preferentially expressed in developing rice inflorescences, which suggests that these phasiRNAs function in early reproductive development in rice (Johnson et al., 2009). In *Arabidopsis*, dcl4, rdr6 and sgs3 mutants that lack phasiRNAs display an accelerated transition from the juvenile to adult phase, indicating that phasiRNAs in *Arabidopsis* are involved in vegetative development (Peragine et al., 2004, Xie et al., 2005). All of these studies imply that phasiRNAs are likely involved in the regulation of the transition from vegetative to reproductive development.

Our earlier *in silico* analyses predicted that citrus miR472, miR482, miR827, miR3954 and miRN20 (also named as miR059) may have the potential to act as triggers of phasiRNAs (Liu et al., 2014b, Wu et al., 2015). A total of 407 potential PHAS loci were systematically predicted in citrus genome (Zheng et al., 2015). In this study, we demonstrate that the 22 nt miR3954 targets a NAC transcript and two citrus-specific non-coding transcripts, triggering the biogenesis of phasiRNAs, which in turn regulates many other NAC homologs.

This article is protected by copyright. All rights reserved.
Overexpression of miR3954 indicated that this pathway might be involved in the induction of early flowering in citrus.

Results

Csi-miR3954 targets one NAC gene and two lncRNAs in citrus

Csi-miR3954, a 22 nt miRNA in Citrus sinensis, was identified by our previous high-throughput sequencing analysis (Xu et al., 2010, Liu et al., 2014b). MIR3954 is a single-copy gene that was located in chromosome 9 of C. sinensis. The hairpin secondary structure of the miR3954 precursor was included in a 148 nt sequence. Within the hairpin structure, a bulge led to the generation of an asymmetric miRNA/miRNA* (22nt/21nt) duplex that yielded a 22 nt mature miR3954 (Figure 1a). RNA gel blots showed that miR3954 was highly expressed in the flowers of all four citrus species, including sweet orange (C. sinensis), pummelo (C. grandis), mandarin (C. reticulata) and Hongkong kumquat (Fortunella hindsii) (Figure 1b). We hypothesized that the higher expression of miR3954 in flower might contribute to regulation of flower development or other reproductive traits. After analyzing the degradome sequencing data, Cs7g22460, Cs1g09600 and Cs1g09635 were identified as target genes of miR3954. Furthermore, 5’-rapid amplification of cDNA ends (5’ RACE) analysis was carried out to detect the cleavage site of each target. As shown in Figure 1c, miR3954-mediated cleavage occurred in the complementary structure of all target genes between the 10th and 11th nucleotide relative to the 5’ end of miR3954, which was consistent with previous researches on miRNA-mediated degradation (Llave et al., 2002, Mallory and Bouché, 2008, Liu et al., 2014a). These data verified that miR3954 targeted the...
Cs7g22460, Cs1g09600 and Cs1g09635 in vivo. After manual annotation, Cs7g22460 was annotated as a NAC transcription factor whereas Cs1g09600 and Cs1g09635 cannot be annotated as protein-coding genes. Then we evaluated the coding potential of Cs1g09600 and Cs1g09635 using an online tool (Coding Potential Calculator, http://cpc.cbi.pku.edu.cn/) (Kong et al., 2007). As a result, the largest predicted open reading frame (ORF) of Cs1g09600 and Cs1g09635 encoded a 29 amino acids peptide (8.64% of the whole transcript) and a 25 amino acids peptide (8.13% of the whole transcript) respectively (Appendix S1). The predicted small peptides could not be detected in our previously published proteomics data (Pan et al., 2009, Pan et al., 2012, Zheng et al., 2014). Additionally, we used Cs1g09600 and Cs1g09635 as query sequences to BLASTN against the nucleotide collection (nr/nt) database, as a result, all significant alignments were from citrus. Considering that Cs1g09600 and Cs1g09635 genes do not have homologous protein-coding genes in other plant species and exhibit a very low confidence of protein-coding potential, we regarded these two genes as citrus-specific non-coding genes.

miR3954 triggers the biogenesis of 21 nt phasiRNAs from target genes

It has been reported that 22 nt miRNAs are capable of triggering the generation of phasiRNAs (Chen et al., 2010). Generally, these phasiRNAs were produced in two forms: 21 nt and 24 nt (Allen et al., 2005, Yoshikawa et al., 2005, Johnson et al., 2009, Zhai et al., 2015). To evaluate whether target genes of miR3954 are potential PHAS loci generating phasiRNAs, we used our previously published sRNA data (Liu et al. 2014b) to calculate the phasing score of all target transcripts according to a previous algorithm (De Paoli et al., 2009, De Paoli et al., 2009, De Paoli et al., 2009).
Xia et al., 2013). Consequently, the phasing scores of all targets were greater than 15 (considered as valid PHAS loci) in all tissues when we set the phase cycle length as 21 nt (Figure 2). The phasing scores of all targets were extremely low when we set the phase cycle length as 24 nt (Figure S1). These analyses indicated that Cs1g09600, Cs1g09635 and Cs7g22460 were all PHAS loci producing 21 nt phasiRNAs in citrus. We computationally obtained all potential 21 nt phasiRNAs produced after the miR3954 cleavage site of each target gene. Twenty-five, twenty and twenty-two phasiRNAs were identified from Cs1g09600, Cs1g09635 and Cs7g22460, respectively (Table S1).

Here, we termed these phasiRNAs following the previously reported nomenclature that identified the polarity and the number of hypothetical DCL processing cycles from the miRNA target site (Allen et al., 2005). For instance, Cs1g09600 3’D1(+) stands for phasiRNAs generated from the sense strand (miR3954-targeted strand) of the 1st duplex that is processed towards the 3’ terminal of Cs1g09600. In general, these phasiRNAs from the same transcript had a similar distribution pattern in different tissues, such as phasiRNAs from the 3rd and 7th site of Cs1g09635, which were the two most abundant phasiRNAs in all tissues (Figure 2 and Table S1). Noticeably, Cs1g09635 had a higher phasing score and lower noise compared to other target transcripts, which indicated Cs1g09635 was a high confident PHAS locus specifically triggered by miR3954 (Figure 2). In total, eight phasiRNAs with an abundant expression level (>100 reads in orange flower) from two lncRNA targets were selected and studied further. They were 3’D1(+), 3’D3(-), 3’D4(-), 3’D6(-) and 3’D8(-) of Cs1g09600 and 3’D1(+), 3’D3(-), 3’D6(-) and 3’D7(+) of Cs1g09635. Particularly, Cs1g09635 3’D6(-) and Cs1g09600 3’D6(-) had the same mature sequence (Table S1).

This article is protected by copyright. All rights reserved.
A tobacco-based transient expression system was used to confirm that miR3954 could trigger the production of phasiRNAs by targeting and cleaving Cs1g09600 and Cs1g09635 since miR3954 is a less-conserved miRNA (it is conserved only in several species, such as peach, apple and potato, and it also called miR7122) and its targets Cs1g09600 and Cs1g09635 are citrus-specific transcripts (i.e. both miR3954 and its non-coding targets are non-conserved in tobacco). In this assay, we used Agrobacterium tumefaciens infiltration to transiently co-overexpress miR3954 and its target transcripts (Cs1g09600 and Cs1g09635) in tobacco leaves. Three days after infiltration, we checked whether the predicted phasiRNAs were produced by RNA gel blot analysis. Leaves of wild plants (WT) and leaves overexpressed only the MIRNA (OX-miR3954) or only the target genes (OX-Cs1g09600/Cs1g09635) were used as negative controls. As a result, in the absence of miR3954, overexpression of Cs1g09600 and Cs1g09635 in tobacco leaves did not generate phasiRNAs (Figure 3). In contrast, all eight phasiRNAs were detected in co-overexpression leaves (Figure 3), demonstrating that miR3954 triggered the formation of phasiRNAs from both Cs1g09600 and Cs1g09635.

**Downstream target genes of phasiRNAs**

To explore the downstream pathways regulated by these phasiRNAs, we identified the target genes of these phasiRNAs across the genome using the degradome data (Liu et al., 2014b). We identified 35 target genes supported by the degradome data in addition to the cis-targeting cognate genes Cs1g09600, Cs1g09635 and their homologs (Table S2). Following criteria previously reported (Addo-Quaye et al., 2008, Addo-Quaye et al., 2009), we divided 35
targets identified from degradome into three categories with different confidence levels. Seven of these targets were grouped into category I (highest confidence level), five were grouped into category II (inferior confidence level) and the rest were grouped into category III (lowest confidence level). The following annotation indicated that these target genes were involved in many aspects of biological processes (Table S2). Interestingly, Cs1g09635 3’D6(-) was predicted to targeted eight NAC domain containing genes, including Cs7g22420, Cs7g22430, Cs7g22450, Cs7g22460, Cs7g22470, Cs7g22540, Cs7g22560 and orange1.1t05093. Most of these NAC genes had a high confidence level (3 belong to category I, 2 belong to category II and 3 belong to category III, see Table S2 and Figure S2) in degradome data and a low penalty for pairing with Cs1g09635 3’D6(-), significantly indicating that Cs1g09635 3’D6(-) might function by targeting NAC transcript factors in sweet orange.

Moreover, 5’ RACE was carried out to confirm whether these NAC transcripts were cleaved at the target site of Cs1g09635 3’D6(-). Because the trigger miRNA of Cs1g09635 3’D6(-), miR3954, was highly expressed in orange flowers, we performed the 5’ RACE analysis in orange flowers. Consequently, our analysis demonstrated that six of these NAC transcripts were cleaved at the target site of Cs1g09635 3’D6(-) in orange (Figure 4). Because the sequences of Cs7g22430 and Cs7g22450 are extremely similar, it is hard to design specific 5’ RACE primers to distinct Cs7g22430 from Cs7g22450. We used the same 5’ RACE primer to identify the cleavage site on Cs7g22430 and Cs7g22450, as a result, we obtained only one Cs1g09635 3’D6(-)-mediated cleavage in Cs7g22430 (Figure 4).
Overexpression of miR3954 promotes the early flowering in citrus

To gain insight into the biological function of miR3954 in citrus, we generated transgenic citrus overexpression lines, in which miR3954 was driven by a constitutively active 35S promoter. Because of the advantage of its short juvenile stage (Zhang et al., 2009), Hongkong kumquat (F. hindsii) was used as an explant for transgenic experiments in this study.

After PCR confirmation with a forward primer in the 35S promoter and a reverse primer in MIR3954, nine independent miR3954-overexpressing lines (OX-miR3954) were validated as transgenic lines (Figure S3). Quantitative RT-PCR (qRT-PCR) and RNA blot hybridization showed that the expression of miR3954 was greatly elevated to varying degrees in all the transgenic lines except OX-miR3954-2 line (Figure 5a). We did not observe any notable phenotypic differences during fruit development in the OX-miR3954 lines compared to WT Hongkong kumquat. Then we sowed seeds from OX-miR3954 and WT Hongkong kumquat in the greenhouse under the same conditions. OX-miR3954 seedlings did not exhibit an observable phenotype in seed germination and vegetative growth compared to WT seedlings. However, the progenies of OX-miR3954-3, OX-miR3954-5 and OX-miR3954-7 blossomed or fruited at seven months after seeding in the greenhouse while the progenies of WT did not show any signs of flowering at that time (Figure 5b). Fourteen months after seeding, all these progenies of OX-miR3954 lines were fruited (Figure 5c). Later, we harvested and sowed more seeds from OX-miR3954-3, -5, -7 and WT, and evaluated the flowering time of each plant by the leaf number at the initial flowering stage. On average, OX-miR3954-3 flowered with 18.5 leaves, OX-miR3954-5 with 19.1 leaves, OX-miR3954-7 with 18.4 leaves, whereas WT flowered with 35.4 leaves (Table S3). These observations indicated that overexpression
of miR3954 significantly shortened the flowering time, accelerating the transition from
vegetative growth to reproductive growth.

**Abundance of miR3954, phasiRNAs and their target genes in the transgenic citrus**

To investigate how the overexpression of miR3954 affect downstream target genes in the
transgenic plants of Hongkong kumquat, the orthologs of *Cs1g09600* and *Cs1g09635* in
Hongkong kumquat were cloned and designated as *Fhi-09600* and *Fhi-09635*. Gene
expression analysis indicated that only *Cs1g09635* was transcribed in a relatively low level
(Figure S4a). Sequence comparison showed that the transcript sequence of *Fhi-09635* was
nearly identical (97% identity) to *Cs1g09635*, completely identical (100%) in the region that
generated phasiRNAs (Figure S4b).

We examined the abundance of miR3954 target genes (*Fhi-09635* and *Fhi-22460*),
phasiRNAs derived from *Fhi-09635* (*Fhi-09635* 3’D1(+), *Fhi-09635* 3’D3(-), *Fhi-09635*
3’D6(-) and *Fhi-09635* 3’D7(+)) as well as NAC targets of *Fhi-09635* 3’D6(-) in the flowers
of OX-miR3954-3, -5, -7 and WT Hongkong kumquat. The expression levels of *Fhi-09635*
and *Fhi-22460* were decreased in all OX-miR3954 lines that were inversely correlated with
the expression pattern of miR3954 (Figure 5a and 6a). The abundance of downstream
phasiRNAs was elevated and ranged from two-fold to seven-fold in different OX-miR3954
lines (Figure 6b and Figure S5). The observation that the elevated level of phasiRNAs
(two-fold to seven-fold, Figure 6b and Figure S5) was not as high as the elevated level of
miR3954 (more than thirtyfold, Figure 5a and 5b) was likely because the miRNA target gene
(*Fhi-09635* or *Cs1g09635*) was not co-overexpressed in our transgenic lines. As mentioned

This article is protected by copyright. All rights reserved.
above, the sequences of Cs7g22430 and Cs7g22450 were extremely similar, we used a pair of universal primers to detect the expression level of Cs7g22430 and Cs7g22450. Consequently, decreased levels of Fhi-22430, Fhi-22450, Fhi-22460 and Fhi-22470 were observed in all OX-miR3954 lines, whereas Fhi-22420, Fhi-22540, Fhi-22560 and Fhi-05093 did not exhibit significant decrease between the OX-miR3954 lines and WT (Figure 6c).

Discussion

To date, many 22 nt miRNAs were reported as triggers of phasiRNAs in diverse plant species and were involved in different regulatory pathways. The regulatory network mediated by miRNA and relevant phasiRNAs can be roughly summarized in Figure 7a. Initially, miRNAs trigger IncRNAs to produce phasiRNAs that turn on the in trans regulation. A classic pathway is the miR390-TAS3 (IncRNA)-tasiRNAs-ARFs, which is widely conserved across a broad range of plants from mosses to seed plants (Williams et al., 2005, Axtell et al., 2006, Heisel et al., 2008, Xia et al., 2017). Additionally, miR2118 and miR2275 target IncRNAs and trigger the biogenesis of 21 nt phasiRNAs and 24 nt phasiRNAs respectively in rice (Johnson et al., 2009, Song et al., 2012). In Arabidopsis, miR173 targets TAS1 and TAS2 (IncRNAs) and triggers the formation of phasiRNAs (Chen et al., 2005, Yoshikawa et al., 2005). miR828-directed tasiRNAs derived from TAS4 (IncRNA) regulate MYBs (Rajagopalan et al., 2006). Alternatively, miRNAs target protein-coding genes as well to generate phasiRNAs that turn on both in cis and in trans regulation (Figure 7a) (Howell et al., 2007, Zhai et al., 2011, Xia et al., 2012, Zhu et al., 2012). For example, the miR828-MYBs pathway was conserved in Rosaceae (Xia et al., 2012, Zhu et al., 2012). PhasiRNAs triggered by
different miRNAs (miR482, miR2118 and miR1507) were generated from nucleotide binding site-leucine-rich repeat genes (NB-LRRs) in Solanaceae, Brassicaceae, Fabaceae and Rosaceae (Zhai et al., 2011, Li et al., 2012, Shivaprasad et al., 2012, Fei et al., 2015, Zhao et al., 2015). PhasiRNA mediating in cis regulation may dramatically amplify the suppressive functions of their trigger miRNAs (Zhai et al., 2011, Xia et al., 2012, Fei et al., 2013).

In this study, our findings revealed a citrus-specific pathway composed of miRNA, phasiRNAs and NAC transcription factors (Figure 7b). Our data demonstrated that the citrus 22 nt miR3954 targeted two lncRNAs (Cs1g09600 and Cs1g09635) as well as a NAC transcript (Cs7g22460). We found that all three target genes generated phasiRNAs from the miR3954-mediated cleavage sites. PhasiRNAs derived from lncRNAs were remarkably abundant, while phasiRNAs derived from the NAC gene had a modest abundance (Table S1). Cs1g09600 and Cs1g09635 may be novel noncoding TAS like (TASL) genes in citrus, because their phasiRNAs target in trans many other functional genes (Table S2). One common phasiRNA, which can be derived from both Cs1g09600 and Cs1g09635 (denoted as Cs1g09635 3’D6(-)), targets in trans eight NAC transcripts including Cs7g22460 (Figure S2).

Thus, we uncovered a potential regulatory network with the involvement of miR3954, its relevant phasiRNAs and NAC genes. In this model (Figure 7b), phasiRNAs derived from Cs7g22460 cannot regulate NAC homologs which may be due to the PHAS locus in Cs7g22460 is not conserved with other NACs (Figure S6). As miR3954 targets Cs7g22460 with no mismatch and targets Cs1g09600 and Cs1g09635 with one mismatch (Figure 1), we hypothesize that when expressed at a low level, miR3954 prefers to target Cs7g22460 and generate phasiRNAs that cannot target additional NAC homologs. On the other hand, when
miR3954 is accumulated at a high level, the other pathway may be initiated by the redundant miR3954, resulting in the production of Cs1g09635 3’D6(-) derived from lncRNAs. Subsequently, Cs1g09635 3’D6(-) elevates the suppression level of NACs by targeting additional NAC homologs (Figure 7b). This hypothesized mechanism provides two different layers in repressing of NAC genes, directly or indirectly, which provides a more flexible and elaborate regulation of NAC transcription factors.

It is intriguing that overexpression of miR3954 leads to an early flowering phenotype in transgenic kumquat. This outcome is consistent to our observation that miR3954 has a specific expression in the flowers of sweet orange, mandarin, pummelo and kumquat (Figure 1b). Unlike the developmental defects in leaves, flowers, and fruits in Arabidopsis caused by transgenic expression of miR156 and miR172 under the control of the CaMV 35S promoter (Glazinska et al., 2009, Wu et al., 2009, Zhu et al., 2009), miR3954 only affects the transition from vegetative growth to reproductive growth in citrus. The resultant phenotype of early flowering could be caused by either direct regulation between miR3954 and NAC gene or by miR3954-lncRNA-phasiRNA-NAC pathway. In Arabidopsis, dcl4, rdr6 and sgs3 mutants that were defective in phasiRNA biogenesis exhibited accelerated vegetative phase transition (Peragine et al., 2004, Xie et al., 2005). This difference may be partially explained that the defects of dcl4, rdr6 and sgs3 in Arabidopsis block the biogenesis of almost all phasiRNAs while the citrus miR3954 is specifically target on NAC genes. Our data reveal a clear clue that the NAC genes are vital targets of both miR3954 and phasiRNAs (Table S2 and Figure 4). As expected, the high accumulation level of miR3954 arouses strong suppression of NAC genes.
(Figure 6). A previous study by overexpression of a NAC-domain protein gene (At2g02450) resulted in a late-flowering phenotype (Yoo et al., 2007). Therefore, it is reasonable to speculate that the suppression of NACs promote early flowering in our case. It is noted that miR3954 is not conserved in Arabidopsis, but exist in many perennial plants (such as citrus, peach and apple), possibly indicates that the pathway miR3954-lcRNA/NAC-phasiRNA-NAC may affect flowering time in perennial plants.

NAC domain containing proteins constitute one of the largest families of transcriptional factors in plants and are involved in many aspects of biological processes, including embryonic, floral and vegetative development, lateral root formation and auxin signaling, plant defense mechanisms and abiotic stress (Olsen et al., 2005). Generally, miR164 targeting of a group of NAC genes is conserved in most plant. However, in Populus trichocarpa, miR6445 also targets NACs that do not overlap with NACs targeted by miR164 (Xie et al., 2016). Similarly, in citrus, NACs targeted by miR3954 and relevant phasiRNAs are different from NACs targeted by miR164 (Figure S7), indicating that NACs regulated by miR3954 may have different functions. In Arabidopsis, expression analyses reveal that NAC genes are involved in flower development and reproduction (Hu et al., 2003, Hennig et al., 2004, Wellmer et al., 2004), and similar results were observed in rice (Kikuchi et al., 2000, Fujita et al., 2010). Moreover, researchers demonstrated that overexpression of a NAC-domain protein gene (At2g02450) results in a late-flowering phenotype (Yoo et al., 2007). All of these previous researches suggested that NAC genes may function as regulator in plant reproductive development. Our work provided clues about the involvement of the
miR3954-phasiRNA-NACs pathway in citrus reproduction development. Characterizing this mechanism fully requires further functional analysis of these NAC gene targeted by phasiRNAs.

In conclusion, our work uncovered an intricate regulatory network consisted of miR3954, two IncRNAs, relevant phasiRNAs and NAC genes. We proposed that the two layers of regulation (directly or indirectly) may provide more flexibility to the regulation of NAC genes by small RNAs. Our data also suggest that the miR3954a mediated regulatory circuit may play a role in the induction of early flowering in citrus.

Experimental procedures

Plant materials

Leaf, flower and fruit of sweet orange (C. sinensis), pummelo (C. grandis), mandarin (C. reticulata) and Hongkong kumquat (F. hindsii) were collected at National Citrus Breeding Center in Huazhong Agriculture University. All samples for RNA isolation were frozen in liquid nitrogen immediately after collection and kept at −80°C until use. The Nicotiana benthamiana used in transient expression analysis system was planted in a growth chamber controlled at 14 h light, 10 h dark, 25°C cycles. The well-developed seeds of F. hindsii were collected at National Citrus Breeding Center in Huazhong Agriculture University. Epicotyls of F. hindsii were harvested through sterilized seeding as reported previously (Tan et al., 2009). Epicotyls of F. hindsii were used as explant sources in citrus transformation.
Bioinformatics analysis

Secondary structure of miR3954 was predicted using RNAfold in the Vienna RNA Package as previously described (Liu et al., 2014b). Dataset and bioinformatics pipeline used in identification of phasiRNAs triggered by miR3954 and their target genes were attainable from previous publication (Liu et al., 2014b). In brief, firstly, miR3954 and its target transcripts were obtained from degradome data. Secondly, all 21 nt or 24 nt standard phasiRNAs generated from the cleavage site of target transcripts to 3’ end of targets were predicted one by one. Finally, we checked the expression level and target transcripts for all predicted phasiRNAs from sRNAome and degradome data respectively. Phasing score was calculated through the algorithm which was developed by De Paoli et al. (2009) and refined in Xia et al. (2013). Phasing score data were shown using Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Gene coding potential was calculated by Coding Potential Calculator (CPC) online (http://cpc.cbi.pku.edu.cn/) (Kong et al., 2007).

Expression analyses of small RNAs and their target genes

To analyze expression levels of miRNAs and phasiRNAs, stem-loop qRT-PCR and RNA gel blot analysis were both used in this study. In both analyses spliceosomal snRNA U6 was used as internal control (Kou et al., 2012). Stem-loop qRT-PCR was performed referring to the previously published method (Chen et al., 2005, Varkonyi-Gasic et al., 2007). Stem-loop RT primers and qPCR primers for miR3954, phasiRNAs and U6 were listed in Table S4. U6 was reversely transcribed by its reverse primer which was also used in the qPCR reaction. Small RNA gel blot analysis was performed as described previously (Liu et al., 2014b). Specific
probes were designed for all sRNAs and U6. The sequences of probes were listed in Table S5. To analyze expression levels of target genes of sRNAs, qRT-PCR was performed where Actin gene was used as a loading control. All target genes were reversely transcribed using random primers/oligo dT primer mix. The primers for target genes used in the qRT-PCR analysis were designed to encompass the cleavage sites mediated by miR3954 or phasiRNAs. The sequences information of primers was listed in Table S6.

Target gene validation

5’ RACE was carried out to verify the cleavage sites of target genes. 5’RACE was performed using GeneRacer Kit (Invitrogen) following the manufacturer’s instructions. Specific outer and inner reverse primers for each target gene used in 5’ RACE analysis were listed in Table S7.

Cloning of MIR3954, Cs1g09600 and Cs1g09635 and construction of overexpression vectors

We designed primers according to the orange genome to clone MIR3954, Cs1g09600 and Cs1g09635. We cloned a 591 nt transcript of MIR3954 containing the 148 nt canonical secondary structure of pre-miR3954 from orange cDNA. Subsequently, the cloned MIR3954 was inserted into an overexpression vector. This OX-miR3954 vector was used in the transient expression in tobacco and citrus transformation. We cloned a 594 nt transcript of Cs1g09600 covering the cleavage site of miR3954 and all aforementioned phasiRNAs. We also cloned a 678 nt transcript of Cs1g09635 covering the cleavage site of miR3954 and all
aforementioned phasiRNAs. These two cloned transcripts were inserted into overexpression vectors respectively. OX-Cs1g09600 and OX-Cs1g09635 were used in transient expression assay. The detailed information of sequences was listed in Appendix S2.

**Transient expression assay in tobacco**

To verify miR3954-mediated cleavage on Cs1g09600 and Cs1g09635 and triggering the biogenesis of downstream phasiRNAs in vivo, we used A. tumefaciens infiltration to co-express miR3954 and its targets in tobacco leaves. Transient expression in tobacco was performed as described previously (Sparkes et al., 2006). A. tumefaciens GV3101 was used as the mediator in this assay.

**Citrus transformation and regeneration**

A. tumefaciens EHA105 was used as the mediator in citrus transformation. The transformation and regeneration procedure were described previously (Duan et al., 2007).

**Cloning orthologous genes from Hongkong kumquat**

As Hongkong kumquat is closely related with Citrus and belongs to the Citrinae group of the subfamily Aurantiodeae in Rutaceae family (Swingle, 1967), primers designed for sweet orange were directly used to clone the orthologs of Cs1g09600, Cs1g09635 and NACs in Hongkong kumquat. After that, we sequenced all the orthologs in Hongkong kumquat. Based on the sequence information, we designed primers of Fhi-09600, Fhi-09635 and Fhi-NACs to perform subsequent qRT-PCR analysis.

This article is protected by copyright. All rights reserved.
Acknowledgements

This work was funded by the National Natural Science Foundation of China (31572105 and 31521092) and the key project of Hubei provincial Natural Science Foundation (2015CFA031). The authors declare no conflict of interest.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Figure S1. 24 nt phasing score distribution along Cs1g09600, Cs1g09635 and Cs7g22460 in orange leaf, flower and fruit tissues.

Figure S2. T-plots of the NAC targets of Cs1g09635 3’D6(-).

Figure S3. Genotyping of miR3954 transgenic lines.

Figure S4. Comparison of non-coding targets of miR3954 in orange and kumquat.

Figure S5. The expression analysis of phasiRNAs from Fhi-09635 in OX-miR3954 lines and WT.

Figure S6. PHAS locus in Cs7g22460.

Figure S7. Comparison of miR3954 and miR164 targeted NACs.

Table S1. Reads of all phasiRNAs derived from Cs1g09600, Cs1g09635 and Cs7g22460 in different orange tissues.

Table S2. Annotation of targets of phasiRNAs derived from Cs1g09600 and Cs1g09635.

Table S3. The flowering times of OX-miR3954 lines and WT kumquat.

Table S4. Primers for qRT-PCR of small RNAs.

Table S5. Probes used in the sRNA gel blot.

This article is protected by copyright. All rights reserved.
Table S6. Primers for qRT-PCR of target genes.

Table S7. Inner and outer primers for 5’ RACE analysis of target genes.

Appendix S1. Coding potential analyses of Cs1g09600 and Cs1g09635.

Appendix S2. Detailed information about sequences used in the construction of overexpression vectors.

Reference:


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


**Figure legends:**

**Figure 1.** Precursor structure, expression pattern and target validation of miR3954. (a) Secondary structure of pre-miR3954 from orange. The sequence of mature miR3954 is highlighted in red, whereas the miR3954* is in green. The red arrow indicates the bulge in the miRNA/miRNA* duplex. (b) The expression level of miR3954 in different tissues of different citrus species. The expression level of miR3954 was analyzed by an RNA gel blot in which snRNA U6 was used as a loading control. L, leaf; F, flower; Fr, Fruit. (c) Target validation was carried out using 5’ RACE analysis. The arrows indicate the cleavage sites in each target and the numbers indicate the frequency of valid sequenced clones.

**Figure 2.** The 21 nt phasing score and sRNA abundance distribution along *Cs1g09600*, *Cs1g09635* and *Cs7g22460* in orange leaf, flower and fruit tissues. miR3954 Phasing scores were evaluated based on the mapping results of 21 nt sRNAs. The length of intervals between dotted lines is 21 nt. The miR3954 mediated cleavage sites of each target transcript are...
indicated by the red dotted lines and the red arrows. Genomic configurations of each target are shown at the bottom of each panel.

**Figure 3.** miR3954 targets *Cs1g09600* and *Cs1g09635* and triggers the biogenesis of phasiRNAs. Tables at the top of the figure indicate the *A. tumefaciens* expression vectors used in transient infiltration. From left to right, each lane contains total RNAs extracted from wild type tobacco leaves, OX-miR3954 leaves, OX-*Cs1g09600/Cs1g09635* leaves, and co-overexpression leaves. The expression of miR3954 and relevant phasiRNAs were analyzed by RNA gel blot. Probes used in each analysis are listed on the left. The snRNA U6 was used as a loading control. The expression of two IncRNAs was analyzed by qRT-PCR, and the *actin* gene was used as a loading control. Data from the qRT-PCR experiments are represented as the mean of four biological replicates (error bars indicate the standard deviation).

**Figure 4.** Target validation of *Cs1g09635 3’D6(-)* in orange flower. 5’ RACE analysis was carried out to verify the NAC targets of *Cs1g09635 3’D6(-)*. Arrows indicate the cleavage sites in each target and the numbers indicate the frequency of valid sequenced clones.

**Figure 5.** Transgenic overexpression of miR3954 accelerates flowering in citrus. (a) The expression level of miR3954 in each transgenic line. The expression level was analyzed by an RNA gel blot and qRT-PCR respectively. snRNA U6 was used as a loading control in both analyses. Data from the qRT-PCR experiments are represented as the mean of three biological replicates (error bars indicate the standard deviation). Early flowering (b) and fruiting (c) of miR3954 overexpression lines in the greenhouse. Left panels show individual progenies of OX-miR3954-3, OX-miR3954-5 or OX-miR3954-7 and the right panels show the WT
kumquat progenies growing in the greenhouse. Seeding dates are also listed on the right side. Scale bars: 1cm.

**Figure 6.** The expression analysis of miR3954 targets, downstream phasiRNAs and NAC targets of Cs1g09635 3’D6(-) in OX-miR3954 lines and WT. The expression level was analyzed by qRT-PCR. snRNA U6 and the *actin* gene were used as loading controls in expression analysis of phasiRNA and target genes respectively. Data from the qRT-PCR experiments are represented as the mean of three biological replicates (error bars indicate the standard deviation). (a) The expression analysis of miR3954 targets in OX-miR3954 lines and WT. (b) The expression analysis of phasiRNAs from *Fhi*-09635 triggered by miR3954 in OX-miR3954 lines and WT. (c) The expression analysis of NAC target genes of Cs1g09635 3’D6(-) in OX-miR3954 lines and WT.

**Figure 7.** Summarization of different regulatory networks among miRNA, relevant phasiRNA and target genes. (a) miRNA cleaves long noncoding RNAs or protein coding genes and triggers the biogenesis of phasiRNAs that regulate downstream target genes. The pathway in which miRNA targets lncRNAs and turns on *in trans* regulation is also termed as the tasiRNA pathway. The pathway that miRNA cleaves protein coding genes and trigger the biogenesis of phasiRNAs amplify the suppressive function of the trigger miRNAs (*in cis*) and *in trans* regulate downstream targets. (b) A two-layer regulatory network of csi-miR3954 in citrus. Layer one is a direct approach (in orange color) in which miR3954 directly regulates a NAC gene (*Cs7g22460*). Layer two is an indirect approach (in green color) in which miR3954 triggers the formation of Cs1g09635 3’D6(-) by targeting two lncRNAs, and then Cs1g09635 3’D6(-) regulates eight NAC homologs including *Cs7g22460*.

This article is protected by copyright. All rights reserved.
This article is protected by copyright. All rights reserved.
This article is protected by copyright. All rights reserved.