Research Article

Novel miR390-Dependent Transacting siRNA Precursors in Plants Revealed by a PCR-Based Experimental Approach and Database Analysis

M. S. Krasnikova,[1] I. A. Milyutina,[2] V. K. Bobrova,[2] L. V. Ozerova,[3] A. V. Troitsky,[2] A. G. Solovyev,[1, 2] and S. Y. Morozov[2]

[1] *Institute of Agricultural Biotechnology, Russian Academy of Agricultural Sciences, Timiryazevskaya 42, 127550 Moscow, Russia*
[2] *A.N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, 119992 Moscow, Russia*
[3] *Main Botanic Garden Russian Academy of Sciences, Botanicheskaya 4, 127276 Moscow, Russia*

Correspondence should be addressed to S. Y. Morozov, morozov@genebee.msu.su

TAS loci in plant genomes encode transacting small interfering RNAs (ta-siRNAs) that regulate expression of a number of genes. The function of TAS3 precursor in *Arabidopsis thaliana* is controlled by two miR390 target sites flanking two ta-siARF sequences targeting mRNAs of ARF transcription factors. Cleavage of the 3′-miR390-site initiates ta-siRNAs biogenesis. Here we describe the new method for identification of plant ta-siRNA precursors based on PCR with oligodeoxyribonucleotide primers mimicking miR390. The method was found to be efficient for dicotiledonous plants, cycads, and mosses. Based on sequences of amplified loci and a database analysis, a novel type of miR390-dependent TAS sequences was identified in dicots. These TAS loci are characterized by a smaller distance between miR390 sites compared to TAS3, a single copy of ta-siARF, and a sequence conservation pattern pointing to the possibility that processing of novel TAS-like locus is initiated by cleavage of the 5′-terminal miR390 target site.

1. Introduction

Small RNA-mediated gene silencing plays important roles in many cellular processes, including development, genome maintenance and integrity, and adaptive responses to biotic and abiotic stress in most eukaryotes. Small RNAs, usually 20–25 nucleotides (nt) in length, guide heterochromatin formation, mRNA degradation, translational inhibition, and DNA elimination. In plants, small RNAs are highly diverse and significant progress has been achieved in unraveling the components and mechanisms involved in their biogenesis and function [1–8]. MicroRNAs (miRNAs) are a class of small noncoding RNAs (21-22-nt long) transcribed from the genomes of all multicellular organisms and some viruses [5, 9–13]. Plant miRNA biogenesis starts with the transcription of miRNA precursors by RNA polymerase II. These precursors contain the mature miRNA sequence within the stem of a long imperfect RNA hairpin which is processed

in the nucleus by RNaseIII-like protein DCL1 to give an imperfect RNA duplex with 2 nt 3′ overhangs on each strand [5, 14, 15]. In total, several dozens of plant miRNA families (hundreds of individual miRNA species) are currently identified and found to target mainly protein-coding mRNAs [7, 12, 13, 15, 16]. However, some microRNAs guide the cleavage of the non-protein-coding primary transcripts of TAS genes directing the formation of trans-acting siRNAs (ta-siRNAs). In this case, miRNAs represent a cleavage guider for ARGONAUTE (AGO) proteins with RNaseH-like activity that cleaves TAS single-stranded RNA transcript in the region complementary to small RNA [4]. TAS cleavage product(s) is converted by cooperative action of RNA-binding protein SGS3 and RNA-polymerase RDR6 to double-stranded form and subsequently processed by DCL4 to produce a cluster of ta-siRNAs that are phased in 21 nt increment relative to the original cleavage site on both strands [4, 5, 13, 17, 18]. The generated 21 nt ta-siRNAs further work as components

of RISC complex to guide AGO-dependent cleavage of their target mRNAs. Recent studies have reported that some ta-siRNAs also work as cleavage guiders to cut other TAS RNA precursors similarly to microRNA and thus generate secondary ta-siRNAs with new specificity [19]. Most TAS RNA precursors have only single miRNA target motif (e.g., miR173) positioned 5′ to ta-siRNA-producing site, and the target motif is cleaved by AGO1 guided by the respective miRNA. However, two miR390 target sites (5′ and 3′ to ta-siRNA site) were shown to be necessary for TAS3 precursor RNA cleavage that appeared to be dependent on specific interaction between AGO7 and miR390 [4, 18, 20].

In this study, we analyzed TAS3-related sequences encoded by genomes of different land plants. Using multiple alignments of TAS3-like RNAs in dicotyledonous plants [17, 18], we synthesized oligonucleotide primers complementary to miR390 target sites positioned 5′ and 3′ to ta-siRNA site. In a control PCR reaction with *Arabidopsis thaliana* chromosomal DNA, we obtained a PCR product of expected size. However, in *Nicotiana benthamiana* and tobacco, in addition to the PCR products corresponding to TAS3-like RNAs, we found smaller PCR fragments. Sequencing of these fragments and further database analysis revealed that they represent a new, previously undescribed, type of miR390-mediated TAS genes producing potential precursors of ta-siARF RNA in different dicotyledonous plants.

2. Materials and Methods

2.1. Plant Material. Plants were taken from collections of Main Botanic Garden of the Russian Academy of Sciences and Department of Virology, Moscow State University. Solanaceous plants were germinated in standard growth chamber conditions for several weeks until stage of 6–8 leaves.

2.2. Analysis of Nucleic Acids. Genomic plant DNA was isolated from 200 mg of plant material by DNA extraction kit (Macherey-Nagel) according to the protocol of the manufacturer. Total RNA was isolated from young tobacco leaves with the Trizol reagent according to the manufacturer's instructions (Invitrogen). Digestion of any contaminating DNA was achieved by treatment of samples with RQI RNase-free DNase (Promega). Reverse transcription was performed with 1 μg of total RNA and oligo (dT)-primer using the RT system (Invitrogen) according to the protocol of the manufacturer. Primers for dicotiledonous plants and *Cycas revoluta* were: forward primer: TAS-P 5′-GGT-GCTATCCTATCTGAGCTT-3′ and reverse primers TAS-Mcaa 5′-AGCTCAGGAGGGATAGCAA-3′ and TAS-Maca 5′-AGCTCAGGAGGGATAGACA-3′.

The primers for mosses (Physcomitrella-specific) were: forward primers: Bryo TAS-P1 5′-GGCGCTATCCCTCCT-GAGCT-3′ and Bryo TAS-P3 5′-GACGCTACCCTTCCT-GAGCT-3′, reverse primer: Bryo TAS-M 5′-TAGCTCAGG-AGTGATA(G/T)A(C/A)AA-3′. For PCR, 25–35 cycles were used for amplification with a melting temperature of 95°C, an annealing temperature of 58°C (for seed plants) or 60°C

(for mosses), and an extending temperature of 72°C, each for 30 seconds, followed by a final extension at 72°C for 3 minutes. PCR products were visualized by electrophoresis of samples in 1% agarose gel. For cloning, the PCR-amplified DNA bands were isolated from gel and ligated into pGEM-T (Promega). In each case, 7–18 independent clones were sequenced. Sequences of 20–50% of clones depending on particular plant species showed no presence of tasi-ARF RNA sequences and these clones were regarded as false ones. In some cases (less than 10% of clones), nucleotide sequences of tasi-ARF precursor genes have the ends (primer sites) which are different by single nucleotide substitutions or deletions from the primer sequences. We propose that this discrepancy may result from rare mutations after cloning of PCR products in *Escherichia coli*.

2.3. GeneBank Accession Numbers. DNA sequences were deposited at NCBI databank under the following accession numbers: *Nicotiana benthamiana*—FJ804742; *Nicotiana tabacum*—FJ804743 ("short" TAS) and FJ804751 (classical TAS); *Datura stramonium*—FJ804744; *Solanum demissum*—FJ804745; *Physalis longifolia*—FJ804746; *Brachythecium latifolium* (clone 50 Br)—FJ804747; *Brachythecium latifolium* (clone 47 Br)—FJ804748; *Hookeria lucens*—FJ804749; *Cycas revolute*—FJ804750.

3. Results

To design primers for PCR and cloning of TAS3-like genes from *Nicotiana benthamiana* and *Nicotiana tabacum*, we used a multiple alignment of TAS3-like RNA precursors from flowering plants published by Axtell and others [18]. PCR analysis was performed with two pairs of primers: P-Tas3 corresponded to the 5′ miR390 target site and either M-Tas3/caa or M-Tas3/aca, both complementary to the 3′ miR390 target/cleavage site [18, 20]. Control PCR reaction using chromosomal *Arabidopsis thaliana* DNA as template and both pairs of primers resulted in efficient synthesis of a single PCR-fragments with expected size of 260 bp (Figure 1(a)) that was in agreement with calculated distance between 5′ and 3′ miR390 target sites in *Arabidopsis* ta-siARF precursor RNAs [17, 18]. Sequencing of these PCR-amplified and cloned DNA fragments confirmed specific amplification of at least two of three known *A. thaliana* TAS3 chromosomal loci (data not shown) [17, 21].

PCR amplification of chromosomal DNA from *Nicotiana benthamiana* and *Nicotiana tabacum* (cv. Samsun) also resulted in synthesis of one major band of 250 bp and 280 bp, respectively (Figure 1(a) and data not shown). Cloning and sequencing of the obtained DNA bands revealed that the amplified sequences contained putative ta-siARF site composed of two tandem copies of ARF-specific ta-siRNAs and located between miR390 target sites corresponding to PCR primers. Moreover, amplified sequences showed obvious similarity to TAS3-like genes from other dicotyledonous plants (Figure 2) [17, 18]. Bioinformatic analysis of the putative TAS3-like sequences from *Nicotiana benthamiana* and *Nicotiana tabacum* using NCBI Blast revealed closely

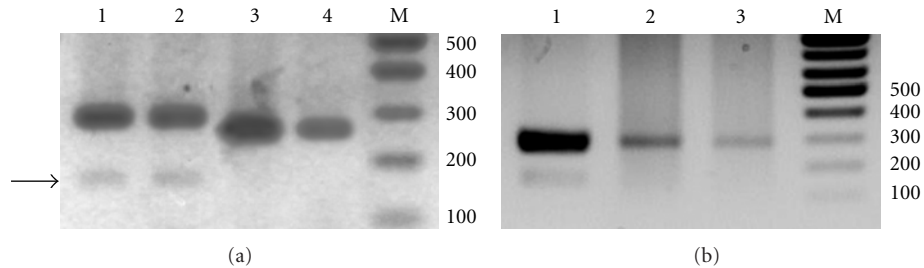(a)   (b)

FIGURE 1: Agarose gel analysis of PCR products. (a) Amplification of genomic DNA sequences flanked by miR390 sites. PCR products were obtained on *Nicotiana tabacum* genomic DNA with primers P-Tas3 and M-Tas3/caa (Lane 1), on *N. tabacum* genomic DNA with primers P-Tas3 and M-Tas3/aca (Lane 2), on *Arabidopsis thaliana* genomic DNA with primers P-Tas3 and M-Tas3/caa (Lane 3), and on *A. thaliana* genomic DNA with primers P-Tas3 and M-Tas3/aca (Lane 4). Arrow points to the 170-nt PCR product. M is DNA size markers. Marker fragment sizes in base pairs are indicated on the left. (b) Amplification of *N. tabacum* genomic DNA and cDNA preparation. PCR products were obtained on *N. tabacum* genomic DNA with primers P-Tas3 and M-Tas3/caa (Lane 1), on *N. tabacum* cDNA with primers P-Tas3 and M-Tas3/caa (Lane 2), and on *N. tabacum* cDNA with primers P-Tas3 and M-Tas3/aca (Lane 3). M is DNA size markers. Marker fragment sizes in base pairs are indicated on the left.

FIGURE 2: Nucleotide sequence alignment of "280-nt-long" ta-siARF loci of dicot plants. Sequences of *Nicotiana tabacum* and *Nicotiana benthamiana* (undelined) were determined in this study, the others were from GeneBank. Boxes indicate miR390 complementary sites and ta-siARF-encoding sequence regions. The length of nonconserved sequence between the 5′-miR390 complementary site and ta-siARF sequence not included in the alignment is indicated. Shading shows residues identical for the given position in most sequences. Sequences of *N. tabacum* and *N. benthamiana* miR390 complementary sites originate from primers used for the amplification of ta-siARF loci. GeneBank accession numbers of the sequences included in the alignment are as follows: *Nicotiana tabacum*, FJ804751; *Nicotiana benthamiana*, FJ804742; *Arabidopsis thaliana*, BX838290; *Glycine max*, BE330988; *Manihot esculenta*, CK652751; *Malus domestica*, CN490861; *Vitis vinifera*, DT025007; *Populus trichocarpa*, DT498974; *Lycopersicon esculentum*, DV105041.

related sequences in representatives of the family *Solanaceae*. For example, *Nicotiana tabacum* (cv. Samsun) TAS3-like sequence showed almost 100% identity to the *Nicotiana tabacum* (cv. Hicks Broadleaf) genomic sequence obtained in the frame of Tobacco Genome Initiative (GenBank accession no. FH434354) and 78% identity to the *Solanum lycopersicum* (cv. Heinz 1706) chromosome 11 BAC clone (GenBank accession no. DU917444) (Figure 2).

Unlike *Arabidopsis thaliana* DNA showing only one visible DNA band after PCR amplification, *Nicotiana tabacum* DNA consistently produced a second minor PCR product of about 170 bp in length (Figure 1(a)). Since no indications on existence of miR390-dependent TAS3-like genes with so closely arranged miRNA target sites were found in the literature and sequence databases, cloning and sequencing of

the minor PCR fragment was carried out. The 170 bp DNA sequence was found to contain the putative ta-siARF site composed of one copy of ARF-specific ta-siRNAs (Figure 3). To rule out the possibility of PCR/cloning artifact we compared this sequence to the available nucleotide sequence databases. As it was found for the 280-nucleotide-long *Nicotiana tabacum* (cv. Samsun) PCR product, 170 bp fragment contained sequence closely related to some yet uncharacterized database entries (Figure 3). Particularly, it showed 96% identity to *Nicotiana tabacum* (cv. Hicks Broadleaf) genomic survey sequences (GSS) obtained in the frame of Tobacco Genome Initiative (GenBank accessions nos. FH734100, FH203124 and FH011695), 91% identity to EST sequence of *Nicotiana tabacum* (cv. SNN) (GenBank accession nos. AM791738), 79% identity to the *Solanum tuberosum* EST
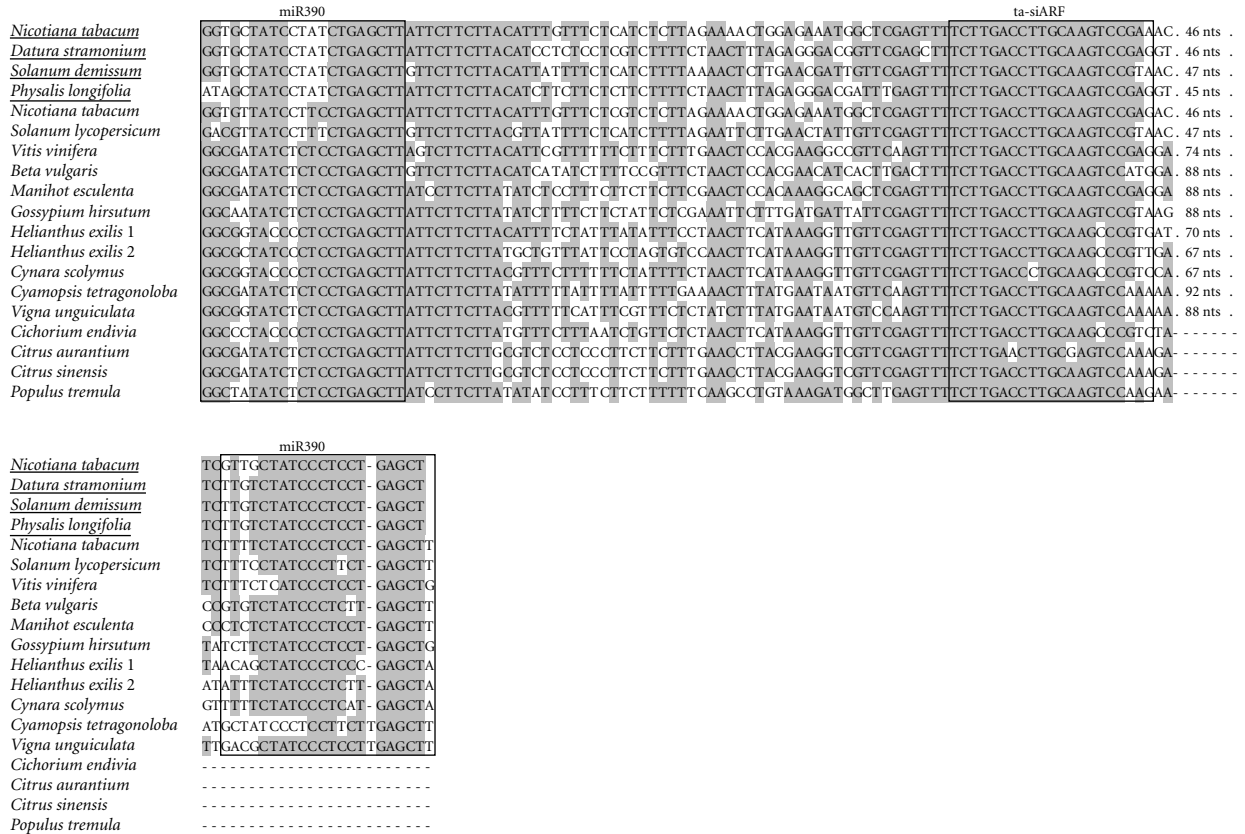
	miR390		ta-siARF	
Nicotiana tabacum	GGTGCTATCCTATCTGAGCTT	ATTCTTCTTACATTTGTTTCTCATCTCTTAGAAAACTGGAGAAATGGCTCGAGTTT	TCTTGACCTTGCAAGTCCGAAAC	. 46 nts .
Datura stramonium	GGTGCTATCCTATCTGAGCTT	ATTCTTCTTACATCCTCCTCGTCTTTTCTAACTTTAGAGGGACGGTTCGAGCTT	TCTTGACCTTGCAAGTCCGAGGT	. 46 nts .
Solanum demissum	GGTGCTATCCTATCTGAGCTT	GTTCTTCTTACATATTTTCTCATCTTTTAAAACTCTTGAACGATTGGTTCGAGTTT	TCTTGACCTTGCAAGTCCGTAAC	. 47 nts .
Physalis longifolia	ATAGCTATCCTATCTGAGCTT	ATTCTTCTTACATCTTCTTCTTCTTCTTTTTCTAACTTTAGAGGGACGATTTGAGTTT	TCTTGACCTTGCAAGTCCGAAAC	. 45 nts .
Nicotiana tabacum	GGTGTTATCCTTCCTGAGCTT	ATTCTTCTTACATTTGTTTCTCGTCTCTTAGAAAACTGGAGAAATGGCTCGAGTTT	TCTTGACCTTGCAAGTCCGAAC	. 46 nts .
Solanum lycopersicum	GACGTTATCCTTTCCTGAGCTT	GTTCTTCTTACAGTTATTTTCTCATCTTTTAGAATTCTTGAACTATTGTTCGAGTTT	TCTTGACCTTGCAAGTCCGTAAC	. 47 nts .
Vitis vinifera	GGCGATATCTCTCCTGAGCTT	AGTCTTCTTACATTCGTTTTTTCTTTCTTTGAAGTCCACGAAGGCCGTTCAAGTTT	TCTTGACCTTGCAAGTCCGAGGA	. 74 nts .
Beta vulgaris	GGCGATATCTCTCCTGAGCTT	GTTCTTCTTACATCATATCTTTTCCGTTTCTAACTCCACGAACATCACTTGACTTTT	TCTTGACCTTGCAAGTCCATGGA	. 88 nts .
Manihot esculenta	GGCGATATCTCTCCTGAGCTT	ATCCTTCTTATATCTCCTTTCTTCTTCTTCGAACTCCACAAAGGCAGCTCGAGTTT	TCTTGACCTTGCAAGTCCGAGGA	. 88 nts .
Gossypium hirsutum	GGCAATATCTCTCCTGAGCTT	ATTCTTCTTATATCTTTTCTTCTTATTCTCGAAATTCTTTGATGATTATTCGAGTTT	TCTTGACCTTGCAAGTCCGTAAG	. 88 nts .
Helianthus exilis 1	GGCGCTACCCCTCCTGAGCTT	ATTCTTCTTACATTTATTTATTTATATTTCCTAACTTCATAAAGGTTGTTCGAGTTT	TCTTGACCTTGCAAGCCCGTGAT	. 70 nts .
Helianthus exilis 2	GGCGCTATCCCTCCTGAGCTT	ATTCTTCTTATGCTGTTTATTCCTAGTGTCCAACTTCATAAAGGTTGTTCGAGTTT	TCTTGACCTTGCAAGCCCGTTGA	. 67 nts .
Cynara scolymus	GGCGGTACCCCTCCTGAGCTT	ATTCTTCTTACGTTTCTTTTTTCTATTTTCTAACTTCATAAAGGTTGTTCGAGTTT	TCTTGACCCTCAAGCCCGTCCA	. 67 nts .
Cyamopsis tetragonoloba	GGCGATATCTCTCCTGAGCTT	ATTCTTCTTATATTTTTATTTTATTTTTGAAAACTTTATGAATAATGTTCAAGTTT	TCTTGACCTTGCAAGTCCAAAAA	. 92 nts .
Vigna unguiculata	GGCGGTATCTCTCCTGAGCTT	ATTCTTCTTACGTTTTTCATTTCGTTTTCTCTATCTTTATGAATAATGTCCAAGTTT	TCTTGACCTTGCAAGTCCAAAAA	. 88 nts .
Cichorium endivia	GGCCCTACCCCTCCTGAGCTT	ATTCTTCTTATGTTTCTTTAATCTGTTCTCTAACTTCATAAAGGTTGTTCGAGTTT	TCTTGACCTTGCAAGCCCGTCTA	- - - - - -
Citrus aurantium	GGCGATATCTCTCCTGAGCTT	ATTCTTCTTGCGTCTCCTCCCTTCTTCTTTGAACCTTACGAAGGTCGTTCGAGTTT	TCTTGAACTTGCGAGTCCAAAGA	- - - - - -
Citrus sinensis	GGCGATATCTCTCCTGAGCTT	ATTCTTCTTGCGTCTCCTCCCTTCTTCTTTGAACCTTACGAAGGTCGTTCGAGTTT	TCTTGACCTTGCAAGTCCAAAGA	- - - - - -
Populus tremula	GGCTATATCTCTCCTGAGCTT	ATCCTTCTTATATATCCTTTCTTCTTCTTTTTTCAAGCCTGTAAAGATGGCTTGAGTTT	TCTTGACCTTGCAAGTCCAAGAA	- - - - - -

	miR390
Nicotiana tabacum	TCGTTGCTATCCCTCCT- GAGCT
Datura stramonium	TCTTGTCTATCCCTCCT- GAGCT
Solanum demissum	TCTTGTCTATCCCTCCT- GAGCT
Physalis longifolia	TCTTGTCTATCCCTCCT- GAGCT
Nicotiana tabacum	TCTTTTCTATCCCTCCT- GAGCTT
Solanum lycopersicum	TCTTTCCTATCCCTTCT- GAGCTT
Vitis vinifera	TCTTTCTCATCCCTCCT- GAGCTG
Beta vulgaris	CCGTGTCTATCCCTCTT- GAGCTT
Manihot esculenta	CCCTCTCTATCCCTCCT- GAGCTT
Gossypium hirsutum	TATCTTCTATCCCTCCT- GAGCTG
Helianthus exilis 1	TAACAGCTATCCCTCCC- GAGCTA
Helianthus exilis 2	ATATTTCTATCCCTCTT- GAGCTA
Cynara scolymus	GTTTTTCTATCCCTCAT- GAGCTA
Cyamopsis tetragonoloba	ATGCTATCCCTCCTTCTTGAGCTT
Vigna unguiculata	TTGACGCTATCCCTCCTTGAGCTT
Cichorium endivia	- -
Citrus aurantium	- -
Citrus sinensis	- -
Populus tremula	- -

FIGURE 3: Nucleotide sequence alignment of "170-nt-long" ta-siARF loci of dicot plants. Sequences of _Nicotiana tabacum, Datura stramonium_, _Solanum demissum,_ and _Physalis longifolia_ (undelined) were determined in this study the others were from GeneBank. Boxes indicate miR390 complementary sites and ta-siARF-encoding sequence regions. The length of nonconserved sequence between the ta-siARF sequence and 3′-miR390 complementary site not included in the alignment is indicated. Dashes in the sequences of _Cichorium endivia_, _Citrus aurantium_, _Citrus sinensis,_ and _Populus tremula_ indicate the sequence region lacking the 3′-miR390 complementary site that could not be aligned with other sequences. Shading shows residues identical for the given position in most sequences. Sequences of _N. tabacum, D. stramonium_, _S. demissum,_ and _P. longifolia_ miR390 complementary sites originate from primers used for the amplification of ta-siARF loci. GeneBank accession numbers of the sequences included in the alignment are as follows: _Nicotiana tabacum_, FJ804743; _Datura stramonium_, FJ804744; _Solanum demissum_, FJ804745; _Physalis longifolia_, FJ804746; _Nicotiana tabacum_, AM791738; _Solanum lycopersicum_, BE459870; _Vitis vinifera_, CU775162; _Beta vulgaris_, CV301446; _Manihot esculenta_, DV447689; _Gossypium hirsutum_, DW502659; _Helianthus exilis_ 1, EE630512; _Helianthus exilis_ 2, EE657417; _Cyamopsis tetragonoloba_, EG977206; _Vigna unguiculata_, FG922881; _Cynara scolymus_, GE588140; _Cichorium endivia_, EL348656; _Citrus aurantium,_ EY848024; _Citrus sinensis_, CK935773; _Populus tremula_, DN500355.

sequence (GenBank accession no. FG548921), 78% identity to the _Solanum lycopersicum_ chromosome 12 clone LE_HBa-26C13 (GenBank accession nos. AC209585) and the _Solanum lycopersicum_ EST sequence (GenBank accession nos. BE459870). Somewhat lesser identity (73-74%) was also revealed for EST sequences from _Solanum phureja_ (GenBank accession nos. FG647537) and _Solanum tuberosum_ (GenBank accessions nos. BQ514736 and BI431636) (Figure 3 and data not shown). These data indicated that 170 bp-long PCR-amplified fragments corresponded to the genuine _Nicotiana tabacum_ (cv. Samsun) genome fragment potentially encoding a new type of ta-siARF RNA precursors. Sequence analysis revealed that, except their length, the 170-nt-long ta-siARF RNA differ from well-characterized 280-nt-long precursors in the structure of ta-siARF-encoding region: while the longer precursors contain two tandem copies of ta-siARF arranged as the conserved 42-nt region, the smaller

precursors described in this paper have only one monomeric 21-nt ta-siARF copy (Figures 2 and 3).

We further experimentally analyzed the occurrence of such putative TAS precursor RNAs in other _Solanaceae_ representatives. Genomic DNAs of _Datura stramonium_, _Solanum demissum_ and _Physalis longifolia_ were taken as templates for PCR with primers P-Tas3 and either M-Tas3/caa or M-Tas3/aca. In all cases, major 280-nucleotide-long and minor 170 bp amplified fragments were easily detected similarly to _Nicotiana tabacum_ (data not shown). Sequencing of the smaller DNA band confirmed the occurrence of novel putative monomeric ta-siARF precursor sequences in different _Solanaceae_ representatives (order _Solanales_) (Figure 3).

The TAS3 loci and miR390 are known to be well conserved among diverse flowering plants [18, 21, 22]. To explore the possibility that the novel monomeric ta-siARF precursor sequences might have emerged in plant evolution
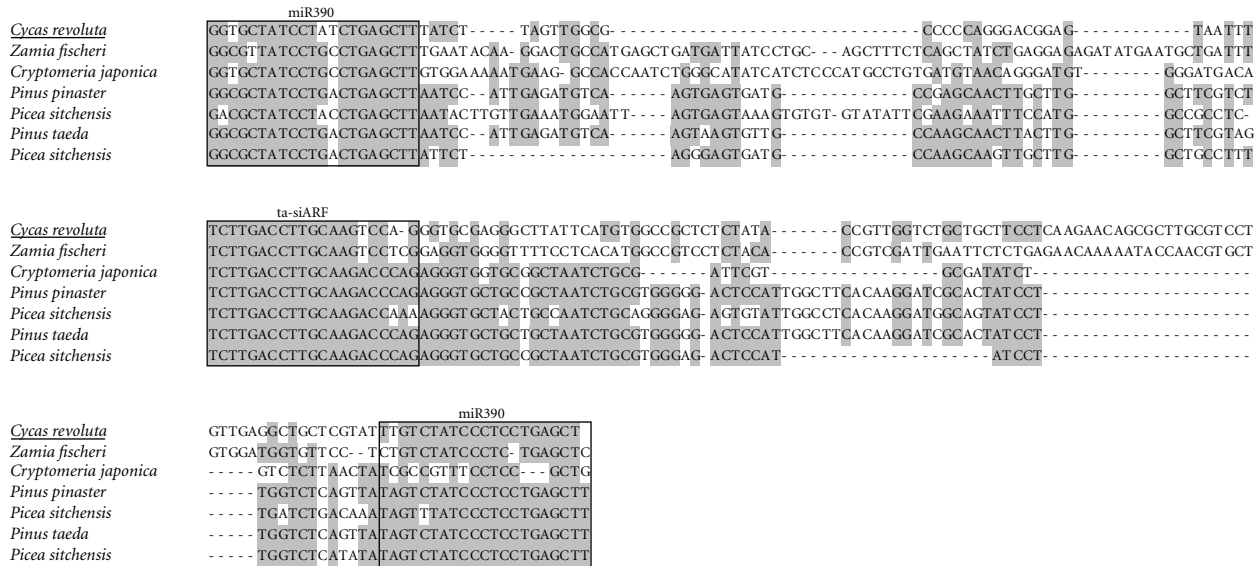
FIGURE 4: Nucleotide sequence alignment of "short" ta-siARF loci of cycad and conifer plants. The sequence of *Cycas revoluta* (undelined) was determined in this study, the others were taken from GeneBank. Boxes indicate miR390 complementary sites and ta-siARF sequence regions. Shading shows residues identical for the given position in most sequences. Sequences of *C. revoluta* miR390 complementary sites originate from primers used for the amplification of ta-siARF loci. GeneBank accession numbers of the sequences included in the alignment are as follows: *Cycas revoluta*, FJ804750; *Picea sitchensis*, EF086492; *Pinus pinaster*, BX682439; *Pinus taeda*, CV034496; *Picea sitchensis*, CO220369; *Zamia fischeri*, DY034932; *Cryptomeria japonica*, DC432705.

earlier than the appearance of family *Solanaceae*, we searched EST database at NCBI. We found that many dicot plants code for putative monomeric ta-siARF loci surrounded in most cases by complementary sites for miR390 positioned 160–190 bp from each other, thus confirming that targeting of ta-siRNA precursors is an evolutionarily conserved function of miR390. These plants include representatives of eurosids—orders *Malpighiales*, *Fabales*, *Malvales,* and *Sapindales*, asterids—orders *Asterales* and *Solanales*, as well as orders *Caryophyllales* and *Vitales* (Figure 3). Analysis of all identified monomeric ta-siARF precursor sequences revealed, in addition to the only copy of ta-siARF sequence, one more characteristic feature. The region between the 5′-miR390 target site and the ta-siARF sequence appeared to be well conserved in length and, to some extent, in sequence, whereas the region between the ta-siARF sequence and the 3′-miR390 target site was found to be variable both in length and sequence (Figure 3). This arrangement of the novel 170 bp-TAS loci differed considerably from that of the well characterized 280 bp-TAS loci, where the sequence and distance conservation was observed for the region between the tandem ta-siARF sequence and the 3′-miR390 target site and not for the sequence between the 5′-miR390 target site and the ta-siARF sequence (Figure 2).

To verify expression of novel putative monomeric ta-siARF precursor RNAs in plant leaves, total RNA was isolated from tobacco plants at 6-8-leaf stage and used as a template for cDNA synthesis with oligo(dT)-primer. It was found that PCR with primers P-Tas3, and either M-Tas3/caa or M-Tas3/aca carried out on this cDNA preparation gave rise only

to 280-nucleotide-long DNA fragment but not detectable 170 bp-fragment (Figure 1(b)).

Conifers (*Pinus taeda*) TAS-like locus, like that of higher plants, was shown to contain the 42-nt ta-siARF region flanked by two miR390 complementary sites. However, unlike TAS3 of flowering plants, only the 5′ miR390-target site is cleaved in conifer TAS3-like RNA precursor [18]. To reveal the possible conservation of TAS3-like loci in lower seed plants, we performed PCR amplification of total DNA from Cycads (*Cycas revoluta*) using dicot-specific primers P-Tas3, M-Tas3/caa and M-Tas3/aca. As it was found for tobacco, PCR on *Cycas revoluta* DNA gave rise to one major DNA fragment of 250 bp and, in addition, several larger and smaller minor bands including those of 170–190 bp in length (data not shown). Sequencing and BLAST analysis of the latter sort of bands revealed that they corresponded to the monomeric ta-siARF precursors of flowering plants (Figure 4). Analysis of databases revealed conifers-encoded sequences showing similarity to the sequenced *Cycas revoluta* putative ta-siARF precursor (Figure 4), showing therefore that this type of regulatory element might have emerged early in seed plant evolution.

Previously, four ta-siRNA loci targeted by miR390 were found in moss *Physcomitrella patens* and referred to as PpTAS3a-d. All four loci contain 5′ and 3′ miR390-target sites, and ta-siRNAs derived from these moss loci regulate, similarly to those encoded by the angiosperm TAS3 loci, ARF genes and, additionally, AP2 transcription factor genes [18, 23, 24]. To confirm applicability of the approach based on PCR amplification of TAS3-like loci to lower plants, we designed new primers BryoTAS-P1,

```
                                                miR390
Brachythecium latifolium 50Br   GACGCTACCCTTCCTGAGCTGAGAATGAAGGCAAGGG--TGGAGGGAGGGCGTGCGGGCTGTGCCGGGCGGCGCCTTGTTAACGGGGTGTTAAGCACCAACGGGC
Physcomitrella patens TAS3d     GGCGCTATCCCTCCTGAGCTTAGAAAGAAGTCAAGGGCCTCCTTGGGAGCAAAGCCAGGTGGGTC-GACGTCGCCTTGTAAGCGGGGTGTTATGCACTTGCGCTT

                                                                          ta-siAP2                                        ta-siARF
Brachythecium latifolium 50Br   GCCCCGGTGGCCTCAGATGTCTGTTTATGGCTCCGTAGGGTGTGATGAGTGCTTTACCCAACACTCATCCTCTACGCAGCCGACCTACCCTTGTGACATGGGCGCA
Physcomitrella patens TAS3d     GACGCTCCGATCTAAGACGTCAGCTGTAGCTCCATGGGGTGTGATGACTGCTTAGCCCGGCACACACTCTCTGCCCGGCCCACCTACCCTTGTGAGATGGGTAGC

                                                          miR390
Brachythecium latifolium 50Br   GTGCTTATCGGCGCAGCCCATGTCAATTGTATATCACTCCTGAGCTA
Physcomitrella patens TAS3d     GCTGATTTGCGCGACTTCCATGTCGGTTTTATATCACTCCTGAGCTA
```

(a)

```
                                    miR390
Brachythecium latifolium 47Br   GACGCTACCCTTCCTGAGCTAAACA----------GATAGAGAAAGGCGT-------------------------------------------GTCGA
Hookeria lucens                 GGCGCTATCCCTCCTGAGCTAAAT-----------CACAGACAGAGTCGT----------------------------------------------
Physcomitrella patens TAS3b     GGCGGTACCCTTCTGAGCTAAAGA----------AAGATGCAGAGGTCTAG---------------------------------------------
Physcomitrella patens TAS3c     GACGCTACCCTTCCTGAGCTAAGTAGGCACGTGGCCCGAGGCAGAGCTGCAGTGCAATCACTCTGAAAATGAGAGTGCATTGAAACTCGTCCCTGCCTTAAGTCT

                                                  ta-siAP2                                           ta-siARF
Brachythecium latifolium 47Br   TGTGGGCGTGAGAGCTAAGGTGTGATGAGTGCTTTAGCATGACATCTCT--TATGCTCAACACGTACTCCAGAGACAAGACCGTCGTCATGACACCAGGGC
Hookeria lucens                 ----GGCGTGTGGGCTAAGGTGTGATGAGTGCTTTAGCAGATCATCTCCCGCGAGCCCGACCACAGACTGTAGTCACACGGCCCGTCTGTTATGACACAAGGGC
Physcomitrella patens TAS3b     ----GGGAGAACTGGCTAGGGTGTGATGAGTGCTTCCGTAAG-CAACTCATG-CGAGCTCAACCACGTACCCTTGAGACAAGACCGTCTGCCATCGCACAAGGGC
Physcomitrella patens TAS3c     AAGGAATAGATTAGGTAGGGTGGAACGATTGCTTAAGCAAGCAAGTTCATA-TTTGCCCAACCAGCTACCCTTGAGACAAGGCTTGTTTGCCTTTGGGCAAGGGC

                                                              miR390
Brachythecium latifolium 47Br   TTTGCACCTTCGTGTAAGCCCTTGTTGTTTTGTCTATCACTCCTGAGCTA
Hookeria lucens                 TTCGCAGCTTCGTGTAAGCCCTTGTTGTTTTGTCTATCACTCCTGAGCTA
Physcomitrella patens TAS3b     AACGCACCATTGCGTGAGCCCTTGTTGTTTTGTCTATCACTCTTGAGCTA
Physcomitrella patens TAS3c     TTTAGAGATATTCCTAAGCCCTTGTTGTTTTGTCTATCACTCTTGAGCTA
```
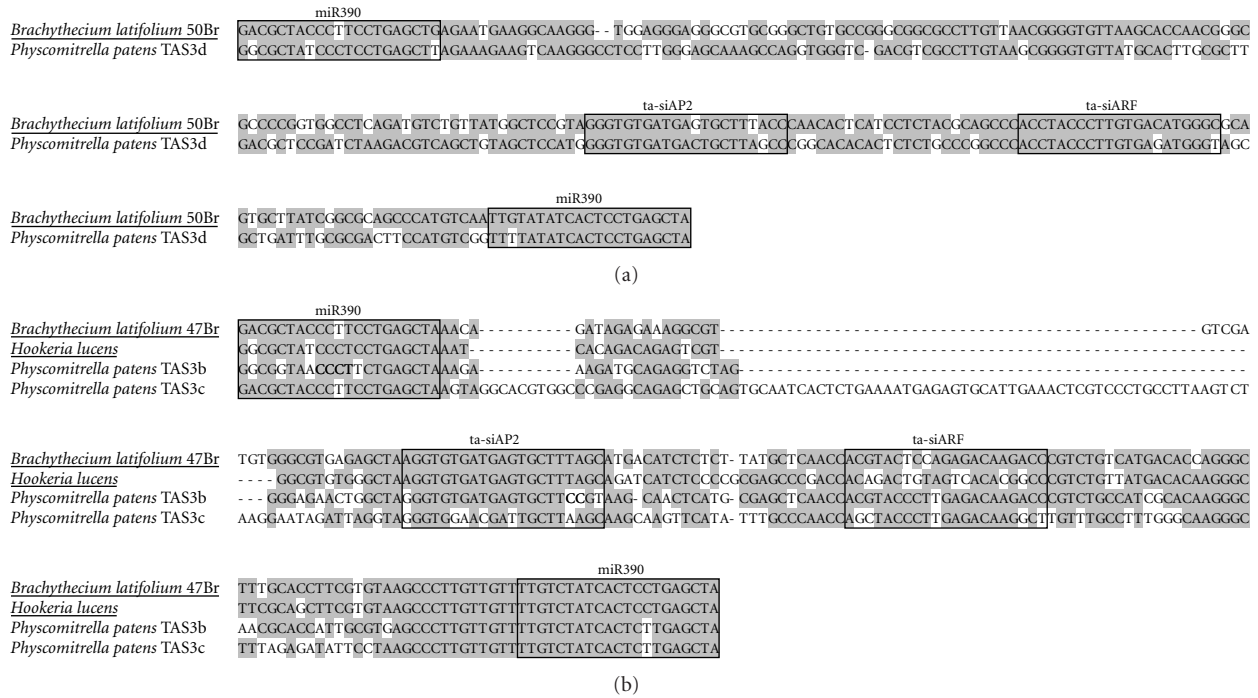
(b)

Figure 5: Nucleotide sequence alignment of ta-siARF loci of mosses. (a) Alignment of *Brachythecium latifolium* clone 50 Br sequence and *Physcomitrella patens* TAS3d. (b) Alignment of *B. latifolium* clone 47 Br sequence and *Hookeria lucens* with *P. patens* TAS3b and TAS3c. Sequences of *B. latifolium* and *H. lucens* (undelined) were determined in this study, the others were from GeneBank. Boxes indicate miR390 complementary sites, ta-siARF sequence regions and ta-siAP2 regions. Shading shows residues identical for the given position in most sequences. Sequences of *B. latifolium* and *H. lucens* miR390 complementary sites originate from primers used for the amplification of ta-siARF loci. GeneBank accession numbers of the sequences included in the alignment are as follows: *Brachythecium latifolium* clone 50 Br, FJ804747; *Brachythecium latifolium* clone 47 Br, FJ804748; *Hookeria lucens*, FJ804749; *Physcomitrella patens* TAS3d, BK005828; *P. patens* TAS3b, BK005826; *P. patens* TAS3c, BK005827.

Bryo TAS-P3, and Bryo TAS-M complementary to miR390-target sites of PpTAS3a-d. These primers were used for PCR on total DNA isolated from *Brachythecium latifolium* and *Hookeria lucens*. Sequencing of major amplified DNA bands showed that *Brachythecium latifolium* encodes at least two types of miR390-dependent AP2/ARF-targeting TAS loci demonstrating sequence similarity to different elements of this kind characterized for *Physcomitrella patens,* namely, PpTAS3d (Figure 5(a)) and PpTAS3b/3c (Figure 5(b)), whereas *Hookeria lucens*, despite its distant taxonomic relation to *Physcomitrella patens,* code for a rather close homolog of PpTAS3b/3c (Figure 5(b)).

4. Discussion

MicroRNAs and ta-siRNAs are small RNAs around 21 nt in length and have widespread roles as posttranscriptional regulators of plant physiology and development [25]. They can affect target genes through a variety of mechanisms, including transcript cleavage and translational repression. The mature miRNAs and ta-siRNAs are derived from larger precursor transcripts that are a few hundred base pairs in size. These precursor RNAs are capped and polyadenylated and can be spliced, in agreement with them being transcribed by RNA polymerase II [12, 26, 27]. Many plant miRNAs and ta-siRNAs are encoded by gene families. For example, four ta-siRNA loci targeted by miR390 were found in *P. patens* and three miR390-dependent TAS loci were identified in *A. thaliana* [17, 18, 21, 23]. Given their similar complementarity to miR390, it was reasoned that multiple TAS3-like loci might be paralogs derived from duplication and divergence of an ancestral locus, and perhaps the areas producing biologically relevant ta-siRNAs may have been preferentially conserved [18, 23]. However, despite obvious progress in studies of miRNA precursor evolution in plants [12, 28, 29], little is known about TAS gene molecular evolution.

Comparing sequences of ta-siRNA precursors within plant genera and between distantly related species should help to determine patterns of molecular evolution that is subject to different evolutionary constraints in comparison with patterns of protein-coding gene evolution. However, the discovery of ta-siRNA genes in plants is still an ongoing process. Although much effort has been directed toward ta-siRNA identification in *Arabidopsis*, poplar and moss *Physcomitrella patens*, many economically or evolutionarily important species have yet to be examined, and more ta-siRNAs likely remain to be discovered even in the species that have been extensively studied. One limiting factor in the ta-siRNA discovery process is the availability of a sequenced

genome, without which a comprehensive analysis of the potential precursor structures of cloned small RNAs is not possible. In *Arabidopsis*, three major methods have been used for ta-siRNA discovery: forward genetics, direct cloning and sequencing, and bioinformatic prediction. Although forward genetic studies have only resulted in the identification of a few ta-siRNAs, this method provides hints to the functions of these miRNAs in addition to their isolation [3]. A second approach to miRNA discovery was bioinformatic prediction. While a number of published studies employed several different algorithms to predict ta-siRNAs, the features that the algorithms search for in the genomic sequences are based on our current knowledge of plant miRNAs and are largely similar among the studies [19]. A third, and perhaps the most effective, method for ta-siRNA discovery was direct cloning and sequencing. In particular, deep sequencing of cloned small RNA libraries using massively parallel signature sequencing (MPSS) or pyrosequencing allowed the identification of numerous small RNAs from *Arabidopsis* and provided a picture of the genomic landscape of small RNAs [18, 23]. One added advantage of this approach is that most small RNA species from a particular genomic loci are exhibited, which helps discern whether the locus gives rise to a single or multiple siRNAs. If multiple small RNAs mapping to both strands of a locus are present, these small RNAs are most likely siRNAs.

Here, we describe a new method for identification of plant miR390-flanked ta-siRNAs precursors based on PCR on the template of genomic DNA or cDNA using oligodeoxyribonucleotide primers mimicking in their specificity miR390. General idea of this experimental approach was suggested by the recently published paper on cloning of animal miRNA-regulated target genes by reverse transcription with direct use of miRNA as a primer [30]. Importantly, the recognition of ta-siRNA precursor by miR390 occurs in most cases by forming imperfect base pairs with 5′ and 3′ target sites [18, 23]. Assuming that imperfect base-pairing is tolerated by Taq polymerase, there is a high chance that designed miR390-related oligodeoxyribonucleotides will recognize both target sites in phylogenetically distant species. To investigate functional constraints on sequence evolution of ta-siARF precursor genes, we have studied sequence variation in the putative precursors of miR390-dependent TAS-like genetic elements between species from the *Solanaceae*. The application of the novel PCR-based method to *Nicotiana tabacum* showed obvious difference from *A. thaliana*: *N. tabacum* DNA directed a major 280-nucleotide-long PCR product and the second minor PCR product of about 170 bp in length (Figure 1). When the method was also applied to *Datura stramonium*, *Solanum demissum,* and *Physalis longifolia,* the minor PCR short product was again identified. Sequencing of 280-nucleotide-long PCR product in the *Solanaceae* and *Arabidopsis* as a control showed that we amplified and cloned the chromosomal region closely related to typical TAS3-related precursor sequences found in diverse dicot plants [17, 18]. This region included 42-nt region from AtTAS3 tasiRNAs that target ARF3 and ARF4 (tasiARFs) and 5′- and 3′- terminal 21-nt regions corresponding to the miR390

complementary sites (PCR primer sites) (Figure 2). These results demonstrate that the PCR-based method is a useful way to clone and identify targets on the basis of base pairing with individual miRNAs. Nucleotide sequence of the minor 170-nucleotide-long product of four selected solanaceous plants revealed the putative 21-nt ta-siARF site composed of one copy of ARF-specific ta-siRNAs and positioned between the miR390 complementary sites (PCR primer sites) (Figure 3). We searched publicly available nucleotide databases of genome survey sequences (GSSs), high-throughput genomics sequences (HTGSs), expressed sequenced tags (ESTs), and nonredundant (NR) nucleotides and identified 16 putative "short" miR390-dependent TAS sequences in 15 diverse plant species (Figure 3 and data not shown).

Using this PCR approach, ta-siARF precursors characterized by the monomeric ta-siARF were found in the cycad *Cycas revolute.* Database search confirmed this experimental finding and revealed novel miR390-dependent TAS sequences in one more cycad (*Zamia fischeri*) and four conifer species (Figure 4). The experimental approach used in this paper was further verified by the analysis of mosses with primers based on *Physcomitrella patens* miR390 complementary; the PCR products obtained on *Brachythecium latifolium* and *Hookeria lucens* DNA represented ta-siRNA precursor sequences that were not known before for these species and related to the previously characterized *P. patens* ta-siRNA loci (Figures 5(a) and 5(b)).

To analyze expression of the novel putative ta-siARF precursor RNAs in plants, total RNA was isolated from tobacco plants at 6-8-leaf stage and used as template for cDNA synthesis with oligo(dT)-primer. PCR analysis of leaf cDNA revealed only 280-nucleotide-long DNA fragment which sequence corresponded to classical tobacco TAS3-like element (Figure 1(b)). On the other hand, RNA template for the EST sequence of *Nicotiana tabacum* (cv. SNN) (GenBank accession no. AM791738) with 91% identity to our "short" TAS sequence from *Nicotiana tabacum* (cv. Samsun) (Figure 3) was isolated from 5-day seedlings. These data suggest that the novel type of miR390-dependent TAS elements might be expressed in tobacco only at early stages of ontogenesis.

Importantly, alignment of the putative novel TAS elements revealed that the length of a region between ta-siARF site and 5′ miR390-target site is highly preserved in contrast to classical TAS3-like genes in which perfect size conservation was revealed between ta-siARF site and 3′ miR390-target site (Figures 2 and 3). In addition, "core consensus" in classical ta-siARF site was found to represent sequence: TCTTGA-CCTTG̲TAAGA̲CCPy, whereas consensus in "short" TAS is different: TCTTGACCTTG̲C̲AAGP̲y̲CCP̲u̲. Significance of this variation remains obscure.

Strikingly, in four cases we failed to reveal the 3′ miR390-target site in "short" miR390-dependent TAS sequences (Figure 3). This type of organization of ta-siRNA precursors is common for the other, except TAS3, TAS genes in *Arabidopsis,* and other plants [4, 12] in which miRNA-dependent cleavage occurs 5′ to ta-siRNA sites. Therefore, we speculate that in "short" miR390-dependent TAS

sequences miRNA-dependent cleavage also can occur 5′ to ta-siRNA sites. This hypothesis requires experimental verification. In addition to identification of the precursor cleavage site, the role of AGO1 and AGO7 in the "short" miR390-dependent ta-siARF precursor processing should be investigated in comparison to classical TAS3-like precursors [4, 20].

Recently a novel silencing method utilizing artificial ta-siRNAs was proved to be efficient for the chimeric *Arabidopsis thaliana* TAS1 and TAS3 genes [31, 32]. This method provides an excellent tool not only for silencing of genes-of-interest but also for investigating functionality of TAS genes and their specific sequence modifications. Application of this method to "short" TAS3-like genes identified in this paper may result in a identification of their possible functions and mapping of cleavage sites in RNA precursors.

Acknowledgments

References

[1] D. Baulcombe, "RNA silencing in plants," *Nature*, vol. 431, no. 7006, pp. 356–363, 2004.

[2] F. Vazquez, "Arabidopsis endogenous small RNAs: highways and byways," *Trends in Plant Science*, vol. 11, no. 9, pp. 460–468, 2006.

[3] H. Vaucheret, "Post-transcriptional small RNA pathways in plants: mechanisms and regulations," *Genes & Development*, vol. 20, no. 7, pp. 759–771, 2006.

[4] H. Vaucheret, "Plant ARGONAUTES," *Trends in Plant Science*, vol. 13, no. 7, pp. 350–358, 2008.

[5] E. Bonnet, Y. van de Peer, and P. Rouzé, "The small RNA world of plants," *New Phytologist*, vol. 171, no. 3, pp. 451–468, 2006.

[6] P. Laporte, F. Merchan, B. B. Amor, et al., "Riboregulators in plant development," *Biochemical Society Transactions*, vol. 35, no. 6, pp. 1638–1642, 2007.

[7] C. Y. Chu and T. M. Rana, "Small RNAs: regulators and guardians of the genome," *Journal of Cellular Physiology*, vol. 213, no. 2, pp. 412–419, 2007.

[8] B. Liu, Z. Chen, X. Song, et al., "*Oryza sativa Dicer-like4* reveals a key role for small interfering RNA silencing in plant development," *The Plant Cell*, vol. 19, no. 9, pp. 2705–2718, 2007.

[9] X. Chen, "MicroRNA metabolism in plants," *Current Topics in Microbiology and Immunology*, vol. 320, pp. 117–136, 2008.

[10] S. Griffiths-Jones, H. K. Saini, S. van Dongen, and A. J. Enright, "miRBase: tools for microRNA genomics," *Nucleic Acids Research*, vol. 36, pp. D154–D158, 2008.

[11] E. Gottwein and B. R. Cullen, "Viral and cellular microRNAs as determinants of viral pathogenesis and immunity," *Cell Host and Microbe*, vol. 3, no. 6, pp. 375–387, 2008.

[12] M. J. Axtell and J. L. Bowman, "Evolution of plant microRNAs and their targets," *Trends in Plant Science*, vol. 13, no. 7, pp. 343–349, 2008.

[13] A. C. Mallory and N. Bouché, "MicroRNA-directed regulation: to cleave or not to cleave," *Trends in Plant Science*, vol. 13, no. 7, pp. 359–367, 2008.

[14] V. Ambros and X. Chen, "The regulation of genes and genomes by small RNAs," *Development*, vol. 134, no. 9, pp. 1635–1641, 2007.

[15] B. Liu, P. Li, X. Li, et al., "Loss of function of OsDCL1 affects microRNA accumulation and causes developmental defects in rice," *Plant Physiology*, vol. 139, no. 1, pp. 296–305, 2005.

[16] A. Takeda, S. Iwasaki, T. Watanabe, et al., "The mechanism selecting the guide strand from small RNA duplexes is different among *Argonaute* proteins," *Plant & Cell Physiology*, vol. 49, no. 4, pp. 493–500, 2008.

[17] E. Allen, Z. Xie, A. M. Gustafson, and J. C. Carrington, "microRNA-directed phasing during trans-acting siRNA biogenesis in plants," *Cell*, vol. 121, no. 2, pp. 207–221, 2005.

[18] M. J. Axtell, C. Jan, R. Rajagopalan, and D. P. Bartel, "A two-hit trigger for siRNA biogenesis in plants," *Cell*, vol. 127, no. 3, pp. 565–577, 2006.

[19] H. M. Chen, Y. H. Li, and S. H. Wu, "Bioinformatic prediction and experimental validation of a microRNA-directed tandem trans-acting siRNA cascade in Arabidopsis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 9, pp. 3318–3323, 2007.

[20] T. A. Montgomery, M. D. Howell, J. T. Cuperus, et al., "Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation," *Cell*, vol. 133, no. 1, pp. 128–141, 2008.

[21] L. Williams, C. C. Carles, K. S. Osmont, and J. C. Fletcher, "A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the Arabidopsis ARF2, ARF3, and ARF4 genes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9703–9708, 2005.

[22] M. J. Axtell and D. P. Bartel, "Antiquity of microRNAs and their targets in land plants," *Plant Cell*, vol. 17, no. 6, pp. 1658–1673, 2005.

[23] M. J. Axtell, J. A. Snyder, and D. P. Bartel, "Common functions for diverse small RNAs of land plants," *Plant Cell*, vol. 19, no. 6, pp. 1750–1769, 2007.

[24] M. D. Howell, N. Fahlgren, E. J. Chapman, J. S. Cumbie, et al., "Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in Arabidopsis reveals dependency on miRNA- and tasiRNA-directed targeting," *Plant Cell*, vol. 19, no. 3, pp. 926–942, 2007.

[25] G. Chuck, H. Candela, and S. Hake, "Big impacts by small RNAs in plant development," *Current Opinion in Plant Biology*, vol. 12, no. 1, pp. 81–86, 2009.

[26] M. Ha, M. Pang, V. Agarwal, and Z. J. Chen, "Interspecies regulation of microRNAs and their targets," *Biochimica et Biophysica Acta*, vol. 1779, no. 11, pp. 735–742, 2008.

[27] M. W. Jones-Rhoades, D. P. Bartel, and B. Bartel, "MicroRNAs and their regulatory roles in plants," *Annual Review of Plant Biology*, vol. 57, pp. 19–53, 2006.

[28] A. Li and L. Mao, "Evolution of plant microRNA gene families," *Cell Research*, vol. 17, no. 3, pp. 212–218, 2007.

[29] N. Warthmann, S. Das, C. Lanz, and D. Weigel, "Comparative analysis of the MIR319a microRNA locus in Arabidopsis and related Brassicaceae," *Molecular Biology and Evolution*, vol. 25, no. 5, pp. 892–902, 2008.

[30] Y. Andachi, "A novel biochemical method to identify target genes of individual microRNAs: identification of a new Caenorhabditis elegans let-7 target," *RNA*, vol. 14, no. 11, pp. 2440–2451, 2008.

[31] T. A. Montgomery, M. D. Howell, J. T. Cuperus, et al., "Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation," *Cell*, vol. 133, no. 1, pp. 128–141, 2008.

[32] M. de la Luz Gutiérrez-Nava, M. J. Aukerman, H. Sakai, S. V. Tingey, and R. W. Williams, "Artificial trans-acting siRNAs confer consistent and effective gene silencing," *Plant Physiology*, vol. 147, no. 2, pp. 543–551, 2008.