Computational estimation and experimental verification of off-target silencing
during post-transcriptional gene silencing in plants

Ping Xu*, Yuanji Zhang*, Li Kang, Marilyn J. Roossinck and Kirankumar S. Mysore
Plant Biology Division, The Samuel Roberts Noble Foundation, Ardmore, OK 73401

Running title: Off-target silencing in plants

* equal contribution

For Submission enquiries contact:
Kirankumar S. Mysore
2510 Sam Noble Parkway, Ardmore, OK 73401
Telephone 580-224-6740; Fax 580-224-6692
e-mail: ksmysore@noble.org

This work was supported by grants from the Samuel Roberts Noble Foundation
ABSTRACT

Successful application of post-transcriptional gene silencing (PTGS) for gene function study in both plants and animals depends on high target specificity and silencing efficiency. By computational analysis with genome and/or transcriptome sequences of 25 plant species, we predicted that about 50-70% of gene transcripts in plants have potential off-targets when used for PTGS that could obscure experimental results. We have developed a publicly available web based computational tool called “siRNA Scan” to identify potential off-targets during PTGS. Some of the potential off-targets obtained from this tool were tested by measuring the amount of off-target transcripts using quantitative PCR. Up to 50% of the predicted off-target genes tested in plants was actually silenced when tested experimentally. Our results suggest that a high risk of off-target gene silencing exists during PTGS in plants. Our siRNA Scan tool is useful to design better constructs for PTGS by minimizing off-target gene silencing in both plants and animals.
INTRODUCTION

Post-transcriptional gene silencing (PTGS), also known as RNA interference (RNAi) in animals, co-suppression in plants, and RNA quelling in fungi, is an epigenetic phenomenon that results in sequence-specific degradation of endogenous mRNAs (Cogoni and Macino, 2000). PTGS is mediated by 21-24 nucleotide (nt) double strand RNA (dsRNA) molecules, termed short/small interference RNAs (siRNAs). The siRNAs are incorporated into a multi-subunit protein complex, the RNAi-induced silencing complex (RISC), which directs the siRNA to and degrades the complementary mRNA (Baulcombe, 2004). Because PTGS allows the silencing of a specific target gene, it has become a popular tool to study gene function in plants and animals (Kamath et al., 2003; Baulcombe, 2004; Hannon and Rossi, 2004; Kuttenkeuler and Boutros, 2004).

In plants, PTGS can be induced by anti-sense and sense transgenic technology, and it is also achieved by expressing dsRNA through stable or transient transformation with RNAi constructs (binary vectors) to knock down the expression of target genes (Varsha Wesley et al., 2001; Miki and Shimamoto, 2004). In addition, a transient PTGS of plant genes by recombinant viruses carrying a near-identical sequence was adapted by a process called virus-induced gene silencing (VIGS) (Baulcombe, 1999; Burch-Smith et al., 2004). Both PTGS approaches, RNAi and VIGS, are becoming powerful tools in functional genomic studies of plants. However, in order to infer gene function through PTGS, it is essential to determine the specificity of mRNAs that are targeted for silencing.

Theoretically, PTGS functions in an siRNA-specific rather than target-specific manner. However, the analyses of mammalian cells transfected with different siRNAs against a target gene by two different research groups led to contradictory conclusions about silencing unintended genes (Chi et al., 2003; Jackson et al., 2003; Jackson and Linsley, 2004). The reasons for these contradictory results are probably due to the differences in experimental designs or microarray analyses. In plants, several RNAi and VIGS studies successfully targeted specific members of gene families for silencing without affecting the transcript level of the most closely related family members or simultaneously silenced a few family members to overcome functional redundancy (Burch-Smith et al., 2004; Hwang and Gelvin, 2004). However, PTGS relies upon
sequences of contiguous nucleotide identity and does not target genes simply based on family relationships. A dsRNA expressed in plants from a binary or virus vector is usually chosen to be identical to a partial or full length sequence of the target gene. Many distinct siRNAs of 21-24 nt can be derived from the cleavage of the dsRNA by Dicer. This may improve the efficiency of RNA silencing, but could also increase the opportunity to suppress unintended genes, the off-targets, containing sequences identical to some of these siRNAs.

Investigating off-target gene silencing is crucial for the accurate interpretation of gene function by PTGS and for the use of PTGS application in agriculture. In the present study, we estimate potential off-targets based on sequence identity for 25 plant species whose genomic or EST sequences are publicly available. Experimentally, we investigated the expression level of several potential off-target genes in an RNAi transgenic Arabidopsis line and in gene silenced (by VIGS) Nicotiana benthamiana plants. In order to assist the design of PTGS constructs to minimize off-target gene silencing or to identify potential off-targets from a particular PTGS construct, we have developed a web-based computational tool. The tool provides an integrated sequence similarity search environment for plant and animal species, identifies potential off-targets, and predicts putatively effective siRNAs from the target query sequences.

RESULTS
Computational analysis of potential off-targets during PTGS.
PTGS in plants is mediated by siRNAs derived from the cleavage of dsRNA produced through binary vector-based RNAi or VIGS constructs. The siRNAs recognize their targets by base pairing. Therefore, the potential off-targets are identified by a direct sequence identity and reverse complimentary sequence identity search. In this study, we define the cDNA producing dsRNA for silencing of the intended target gene (from which the cDNA comes) as the ‘trigger’, and the unintended genes that share a contiguous $\geq 21$ nt region of identity or reverse complementary identity to the trigger as ‘off-targets’. The continuous identical or reverse complementary sequence to a trigger region in an off-target is defined as an ‘off-target region’.
Using the full length cDNA sequence of every transcript in the *Arabidopsis* genome mRNA dataset (ATH1) as a trigger, we estimated the patterns of potential off-targets by searching for identical or reverse complementary regions of $\geq 21$ nt between a trigger sequence and all the other genes except the target itself. Differently spliced transcripts from the same locus in the genome were not counted as additional off-targets. The potential off-target region was scanned along the full length of all released transcript sequences. Off-target regions can be in the 5' and 3' untranslated regions (UTRs) or coding regions because all these regions have been shown to be targeted by siRNAs (McManus et al., 2002; Yu et al., 2002; Yokota et al., 2003; Shirane et al., 2004). By using results obtained from the *Arabidopsis* data analyses, we compared the distributions of off-targets/trigger, off-target sites/off-target, and length of off-target site/off-target site (Fig. 1) with their statistical averages (Table 1). The histograms shown in Fig. 1 have skewed distributions, with majority of the trigger genes having similar potential off-targets as the average (Fig. 1A), majority of the off-targets having approximately same number of off-target sites as the average (Fig. 1B), and majority of off-target sites having almost the same length as the average (Fig. 1C). The average values for off-targets, off-target sites, and lengths of off-target sites are also close to the peak values (medians).

On an average, about 68.7% of *Arabidopsis* transcripts can potentially silence 3.9 off-target genes (Table 1). Each candidate off-target has an average of 3 off-target regions with an average length of over 40 nt (Table 1). However, this could be an overestimation, because not all siRNAs derived from the cleavage of dsRNAs are efficient at silencing (Khvorova et al., 2003; Schwarz et al., 2003). Since little is known about the efficacy of specific siRNAs in plants, we adapted the rules developed by Ui-Tei et al. (Ui-Tei et al., 2004) to predict whether and how many of the trigger sequences mentioned above may produce efficient siRNAs to silence their potential off-targets. Specifically, the rules for 21 nt siRNA duplex with 2 nt overhang include: (i) antisense strand starts with A/U, (ii) sense strand starts with G/C, (iii) at least five of the first seven residues at the 5' terminal of the antisense strand should be (A/U)s, and (iv) there is no G/C stretch of more than 9 nt in length (Ui-Tei et al., 2004). As shown in Table 1, 68.7% of all trigger sequences in the ATH1 database are predicted to generate an average of 35 efficient 21-nt siRNAs to cause potential off-target silencing.
In addition to Arabidopsis, we also estimated the potential off-targets in 24 other plant species, using Gene Indices assembled from Expressed Sequence Tags (ESTs) and other expressed transcripts provided by TIGR (Quackenbush et al., 2001). We first compared the results from Arabidopsis Gene Index version 11 (AGI 11) with the results from the genome mRNAs described above (ATH1). Slightly more trigger sequences were predicted to have potential off-targets when the AGI 11 dataset was analyzed (Table 1). This is because numerous unique gene transcripts in AGI 11 are not represented in the current genome sequence. Estimations of the off-targets in other plant species also indicated high likelihood of off-target silencing (Table 1). The potential off-target effects for plants like rye, pepper and Lotus may be under-estimated in this analysis because of limited sequence availability. For example, the percentage of target sequences predicted to silence off-target is ~ 26% for these species. When those species with < 20,000 EST sequences in their gene indexes are excluded, 50 - 70% of all gene transcripts in each of 17 plant species used as silencing triggers can cause potential off-target silencing. The average numbers of off-targets for each trigger sequence ranges from 2.4 in tomato to 24 in lettuce and sugarcane (Table 1).

Off-target silencing however, may be desirable in PTGS applications to generate a “loss-of-function” phenotype if the target and off-targets are functionally redundant. For example, the same family members or the genes encoding proteins that share conserved functional protein domains may be simultaneously silenced. Arabidopsis gene family information at the Arabidopsis information resource (TAIR) website (http://www.arabidopsis.org/) was used to evaluate the chance that potential off-targets belong to the same gene family as the target gene. The data from TAIR ftp site, after removing ambiguous entries and single-member families, contain 5,842 genes from 674 families (see Supplementary Table S1) with family size ranging from 2 to 307 members. Our analysis indicates that 4,677 gene sequences (79.9% of 5842 genes) of these families may trigger off-target silencing of other members of the same family when a full-length sequence is used as a trigger. Among the families with different numbers of gene members, the number of off-targets was different but generally fewer than six (Fig. 2; Supplementary Table S1). Moreover, only 418 genes from 192 families containing fewer than 10 members, share an identical region of ≥ 21nt with all other members in the same
gene family, so only these transcripts used as trigger sequences can silence all family members. No single trigger sequence from any family with more than 10 members was predicted to silence all members of the family. In addition, the percentage of gene members to be silenced decreases from ~ 79% to < 9% as the family size increases from 2 to > 35. The family CYP705A, in the P450 super-family, is an exception. This family has 24 gene members, and >75% of them were predicted to be potential off-targets when any one of its members is used as RNAi trigger (Fig. 2).

**An integrated search environment for PTGS**

An online tool available at [http://bioinfo2.noble.org/RNAiScan.htm](http://bioinfo2.noble.org/RNAiScan.htm) has been developed to evaluate and minimize the risk of off-target effects during PTGS. The tool is a search environment with several integrated components, including sequence similarity search to identify potential off-targets, efficiency estimation of siRNAs, and functional analysis of off-target genes. In the sequence similarity search, the query sequence is compared against a gene transcript dataset using BLASTn program (Altschul et al., 1997) with user-specified mismatches, which may be tolerable in siRNA-target recognition (Saxena et al., 2003). The datasets include *Arabidopsis* and rice genome mRNAs and all TIGR’s gene indices for plants and animals, and they are updated regularly. All possible siRNAs derived from the query are then evaluated for efficiency using Ui-Tei’s rule (Ui-Tei et al., 2004). The predicted off-target genes can be subject to further functional analysis by identifying Pfam domains (Sonnhammer et al., 1998) and associated gene ontology assignments. The tool can also be used to estimate off-targets during siRNA-induced gene silencing in animals. In addition to displaying the potential siRNAs and off-target candidates, the results page also displays all the potential siRNAs with no sequence similarity to any other genes in the searched database. Therefore the RNAiScan tool can be used to design siRNAs with minimal off-targets.

**Experimental verification of off-target silencing in plants**

The expression of at least 3 off-target genes was knocked down in BT11 RNAi transgenic lines. BT11 is an *Arabidopsis* protein that interacts with the *Agrobacterium* VirB2 protein (Hwang and Gelvin, 2004). A functional study of BT11 through RNAi transgenic
Arabidopsis lines has been published, where dsRNA covering the coding region of *BTII* was produced from the RNAi vector pFGC5941 (Hwang and Gelvin, 2004). We used the coding region of *BTII* as a trigger sequence to search for potential off-targets against the TIGR AGI 12.1 database with the web-based tool "siRNA Scan" described above. We further confirmed the annotations of these off-targets by blast search in NCBI. In addition to one family member, *BTI2*, thirteen other genes were found that share at least 21 nt continuous direct identity or reverse complementary identity to the *BTII* coding sequence (Table 2). These fourteen genes have the potential to be targeted by the siRNAs derived from *BTII*-derived dsRNA. Twelve of them were selected for the analysis of their expression levels in *BTII* RNAi transgenic lines (Table 2). They are TC251703 (*BTI2*), TC251496 (putative 3-isopropylmalate dehydratase large subunit), TC256637 (encodes a probable serine/threonine kinase), TC258543 (RTNLB6), TC255665 (encodes a zinc finger motif protein), TC262843 (encodes a protein of unknown function), TC263798 (encodes a protein of unknown function), TC269146 (encodes a hypothetical protein), TC275407 (encodes a putative serine threonine kinase), TC265975 (encodes a probable membrane protein); TC275528 (encodes a maturase-related protein) and TC275625 (encodes protein of unknown function). In addition, the expression of other 22 genes that share almost 21 nt or 22 nt continuous identity to *BTII* coding region but with one mismatch was also analyzed. The expression levels of these genes as well as the target gene *BTII* were measured by real time PCR and compared with their expression in non-transgenic plants. The *Arabidopsis* elongation factor-alpha (*EF-1*α) gene was used as the endogenous control to normalize the relative transcripts in the reactions. The results showed about 90% down-regulation of target gene *BTII* in one RNAi transgenic line (Fig. 3). Down-regulation of some of the selected potential off-targets, TC262843 and TC269146, and a family member, TC251703 (*BTI2*), was detected in the same RNAi transgenic line. They share at least 23 contiguous nucleotides identity with *BTII* coding region (Table 2). About 75% and 50% down-regulation occurred to TC262843 and TC251703 respectively (Fig. 3). Surprisingly, the expression of TC269146 is about 95% down-regulated, a little more than that of the target gene. TC262843 and TC269146 share the same identical 23nts with *BTII* gene. An oligo DNA with this 23nts-sequence was synthesized and labeled for detection of the potential siRNA involved, but it wasn't
detectable. Nevertheless, a low amount of siRNA could be detected when a 200bp DNA probe of \textit{BTII} containing this 23nts-sequence was used for hybridization (data not shown). No significant difference was seen in the expression of other analyzed genes with at least 21 nts identity or reverse complementarity to \textit{BTII} while very low (unquantifiable) amounts of transcripts were detected for TC255665 in all the tested leaf tissues (data not shown). Also no significant differences were seen in the expression of the selected 22 genes that share almost 21nt or 22nt continuous identity to \textit{BTII} coding region but with one mismatch (Supplementary Figure S1). Similar results were seen from two other independent \textit{BTII} RNAi transgenic lines tested (data not shown).

To rule out the possibility that the lowered expression of some of these selected genes might be due to the loss-of-function of the \textit{BTII} protein instead of off-target silencing, we analyzed the expression levels of all the above investigated genes in an \textit{Arabidopsis} \textit{bti1-2} null mutant line (T-DNA knockout line; Salk-032220). As expected, the expression of \textit{BTII} was below the detectable levels in the mutant plants, but the expression levels of all the other genes tested were similar to those in non-transgenic \textit{Arabidopsis} plants (Fig. 3). The expression of two additional house-keeping genes, \textit{eIF4} and \textit{actin II}, was also investigated for the confirmation of equal RNA amounts used for the real time PCR. Expression levels were similar among the samples from the three groups, further confirming the equal loading of total RNA in the reactions (Fig. 3). Thus, it is off-target silencing caused by \textit{BTII} derived dsRNA that down-regulated the expression of three of the five selected potential off-targets in \textit{BTII} RNAi transgenic plants. In addition, there is a striking phenotypic difference between \textit{BTII} RNAi transgenic lines and \textit{bti1-2} mutant line. Plants from four independent \textit{BTII} RNAi transgenic lines flowered earlier than the wild-type and \textit{bti1-2} mutant plants (Fig. 4) and this is probably due to off-target gene silencing in the RNAi plants.

The expression of some potential off-targets was down-regulated in \textit{N. benthamiana} plants infiltrated with VIGS constructs. Computational analysis showed that more than half of the EST contigs (TCs) in the \textit{N. benthamiana} database have multiple potential off-targets during PTGS (Table 1). We randomly chose two genes for experimental analyses. They are TC381 and TC1146 encoding the U2 small nuclear ribonucleoprotein A’
(snRNAP A') and pyruvate decarboxylase, respectively. Partial fragments from both genes were separately cloned into the Tobacco rattle virus (TRV) RNA2 vector (Liu et al., 2002) for VIGS analysis (their sequences can be found in Supplementary Fig. S2 online). The search with "siRNA Scan" using both fragments as trigger sequences, revealed several potential off-targets (Table 3). Clone TC381 shares contiguous identity of at least 21 nt identity to the following unique sequence entries in the current N. benthamiana database: CK286172, CK288691, CK289650, TC10748, and TC7796. These genes encode proteins that are unrelated to snRNAP A'. Three of these potential off-targets, CK286172, CK288691 and TC10748, were selected for expression analysis by real time PCR in TC381 VIGS-silenced plants. Expression levels of these genes as well as that of the target gene TC381 were compared with their expression in the plants infiltrated with TRV1 (RNA1) + TRV2::00 (RNA2 empty vector). The EF-1α gene was used as the endogenous reaction control to normalize the relative quantity and the expression of the β-tubulin gene was investigated to confirm the equal loading of total RNA. The results showed a greater silencing of TC10748 than the target gene TC381, and very little silencing (not statistically significant) for CK286172 and CK288691 genes (Fig. 5A). The analysis was repeated twice with two different groups of plants and similar results were observed.

A similar scenario was found in TRV2::TC1146 infiltrated plants. Five unique EST entries share at least 21 nt of contiguous reverse complementary sequences to the cloned TC1146 fragment (Table 3). They are CK282591 encoding a protein similar to the VPI/ABI3 family regulatory protein, CK292351 encoding serine carboxypeptidase II-2 precursor and CK287535, CK296810 and TC8666 encoding three different members of the auxin efflux protein family. The expression levels of these genes and TC1146 in silenced plants were compared between the silenced plants and TRV-RNA1 + TRV2:00 infiltrated plants. The expression of target gene TC1146 was decreased to about 50% in TC1146 silenced plants, whereas among the five analyzed potential off-targets, CK287535, CK296810 and TC8666 were down-regulated to a greater extent than TC1146 (Fig. 5B). These three genes all belong to the auxin efflux protein family and have the same reverse complementary sequences of 22 nts to TC1146. The expression level of CK282591 was slightly reduced while there was no significant change in the
expression of CK292351 in TC1146 silenced plants. These experiments were repeated twice with two individual silenced plants and the results were similar. In conclusion, some of the potential off-target genes were silenced to different degrees when VIGS was used to silence target genes of interest in *N. benthamiana*. Because we do not have null mutations of these target genes in *N. benthamiana*, we cannot rule out the possibility that the reduced transcripts of the off-target genes result from the reduction in the amount of target protein.

**DISCUSSION**

The use of RNA silencing/interference for suppressing gene expression has become a powerful and promising approach in gene function exploration and disease treatment in both plants and animals. Its successful application relies on specific and efficient silencing of particular genes or gene families. Exquisite specificity of RNAi through siRNA in animal cells has been supported by several studies (Tuschl et al., 1999; Elbashir et al., 2001; Chi et al., 2003; Semizarov et al., 2003). However, some contradictory reports indicated that siRNA used for RNA silencing can cause off-target suppression at both transcriptional and translational levels in animals. For example, siRNA causes silencing of unintended genes that lack complete sequence identity and sometimes induces non-specific interferon responses in animals (Holen et al., 2002; Amarzguioui et al., 2003; Bridge et al., 2003; Saxena et al., 2003; Sledz et al., 2003; Jackson and Linsley, 2004; Scacheri et al., 2004). Recently, a computational study using the genome and transcriptome sequence data of *Homo sapiens*, *Caenohabditis elegans* and *Schizosaccharomyces pombe* suggested that the risk of transcriptional off-target silencing by siRNA is considerable in all these organisms (Qiu et al., 2005). However, to date there is no experimental evidence in plants to show that off-target silencing of unrelated genes can occur as a result of nucleotide sequence similarity with siRNA.

Direct introduction of siRNA into plant protoplasts for PTGS was reported but it is difficult to be applied in intact plant tissues (Vanitharani et al., 2003). The functions of various plant genes revealed by PTGS have been studied mostly through VIGS or by generating stable transgenic lines that express anti-sense RNA or dsRNA that is identical or complementary to the partial or full-length sequences of target genes. Cleavage of the
expressed dsRNAs by Dicer produces many siRNAs. Because siRNAs of 21-26 nts have been reported in plants (Hamilton et al., 2002; Llave et al., 2002; Tang et al., 2003; Qi et al., 2005), theoretically, 21 nt identity or reverse complementarity between a trigger sequence and a target could be the minimum requirement for successful RNAi. Our computational analyses indicate a high risk of off-target gene silencing among different plant species (Table 1). Results from our experimental analyses with both an RNAi transgenic line and VIGS further confirmed the silencing of some potential off-targets that share at least 22 nts of direct identity or reverse complementary identity to the trigger sequences (Figs. 2 and 4). We were not able to detect the particular siRNAs that have caused off-target silencing through normal northern blot analysis. This may be due to low abundance or rapid degradation of one particular siRNA molecule. Nevertheless, the obvious down-regulation of some of predicted off-targets only occurred in BTI RNAi transgenic lines but not in bti1 mutant plants is compelling. In other studies, a stretch ≥23 nt of perfect sequence identity was found necessary to silence a GFP transgene (Thomas et al., 2001) and heterologous silencing occurred when at least 23 or 24 nt identity existed between the RNAi trigger sequence and the intended gene (Ekengren et al., 2003; Liu et al., 2004). In our analysis, 22nt sequence identity was sufficient to cause off-target silencing. None of the investigated genes with 21 or 22 nts identity but containing one mismatch to the trigger sequence was silenced. Thus at least 22 nts identity may be required for off-target silencing to occur. However, we only analyzed 19 out of 70 predicted off-target candidates with 22 nts identity but containing one mismatch identified by SiRNA scan. Hopefully, a systematic experimental analysis to determine the minimum sequence identity for PTGS between trigger and target sequences will help set a baseline for potential off-target searching with a particular trigger sequence in the future.

Statistically, longer siRNAs should be less likely to silence unintended genes by chance. Therefore predicted off-target risk might be overestimated for siRNAs longer than 21nt. However, because on average each predicted off-target for all plant species analyzed here has multiple off-target regions and each average off-target region is longer than the longest siRNAs reported in plants (Table 1), the overall trend of off-target risk predicted from this study should be valid for longer siRNAs such as 22-26 nts. More
importantly, because siRNAs of various lengths simultaneously coexist within the plant cell (Hamilton et al., 2002), off-target risk should be estimated with the shortest siRNAs. In addition, mismatches in siRNA-target recognition are tolerable and these siRNAs can cause RNA degradation and translational repression in animals (Saxena et al., 2003). It is not known if this applies to plants. In the case of BTI1 RNAi transgenic lines, the transcript accumulation of all the investigated potential off-targets with contiguous at least 22nts identity or with 21 or 22nts identity but containing one mismatch to the expressed dsRNA wasn't affected (Supplementary Fig. S1). However, this is just one particular case. If mismatches of siRNAs can be tolerable in plant PTGS, our computational analysis results may represent the lower-limit of off-target effects as only 21 nt identical regions were counted as potential off-target sites in the present study. Nevertheless, our "siRNA scan" tool includes the option of searching the potential off-targets with complete sequence identity or reverse complementary identity of 18-29 nts to the trigger sequence as well as allowing a few mismatches to the potential siRNAs.

Although our computational analysis showed a high risk of off-target gene silencing during PTGS in plants, the efficiency of off-target gene silencing should be the main factor that affects the functional analysis of a particular target gene. Results from our experimental analyses showed a varied reduction of expression level of the potential off-targets in both Arabidopsis and N. benthamiana. The reduction ranges from none to a greater reduction than that of the target gene expression. The underlying mechanisms for this variation in expression levels of off-targets are not yet clear. Gene silencing efficiency is correlated with siRNA sequence-specific features (Khvorova et al., 2003; Schwarz et al., 2003; Amarzquioui and Prydz, 2004; Reynolds et al., 2004; Ui-Tei et al., 2004), location of the complementary sequence of siRNA in the target (Birmingham et al., 2006), and target accessibility (Luo and Chang, 2004; Pancoska et al., 2004; Brown et al., 2005). At present, target accessibility cannot be reliably predicted as sequences may be partial for most datasets which are assembled from ESTs and other transcripts. Besides, the in vivo mRNA secondary structure is unknown for the datasets in our analysis. Therefore, we only used siRNA sequence properties to assess gene silencing efficiency. Specifically, siRNA efficacy was predicted according to the rules by Ui-Tei et al. (Ui-Tei et al., 2004). A recent comparison study indicates that, among currently
available siRNA efficacy predictors, Ui-Tei’s rules are stable and high performance (Saetrom and Snove, 2004). Although Ui-Tei’s rules have not been experimentally tested in plants, they agree with the requirement for thermodynamic features of efficient siRNA duplex which favors siRNA unwinding by helicases. However, the siRNAs predicted to cause off-target gene silencing had different efficiencies in silencing the target and off-target genes in our experimental analysis. This could be due to the different accessibility of the mRNAs to the siRNA. Earlier, it was shown that siRNAs against different regions of genes display marked variation in their potency in mediating mRNA degradation (Thomas et al., 2001). In addition, strong RNAi effects, as assessed by a phenotypic analysis, were found to correlate with high expression levels of the targeted genes or higher expression of RNAi trigger sequences in C. elegans and plants (Chuang and Meyerowitz, 2000; Cutter et al., 2003; Hu et al., 2004; Kerschen et al., 2004). The structure of siRNA may also have an effect on its efficiency to cause gene silencing. These complex factors together may result in different silencing efficiencies among target and off-target genes.

Short reverse complementary sequences (over 21nt) were found in some completely unrelated genes in our computational analyses. In animals, siRNAs can simultaneously induce sequence-specific degradation of two endogenous mammalian transcripts oriented in opposite directions (Hu et al., 2004). Our experimental results indicate that either strand of the siRNA duplex may silence the genes that contain identical sequences. It is not known whether these endogenous short reverse complementary sequences in different genes have any biological function.

CONCLUSION

The understanding of off-target silencing is crucial for accurate interpretation of gene function by PTGS. Our computational analysis with the genome and transcriptome sequences from 25 plant species showed a high risk of off-target gene silencing when a full length sequence of each transcript entry in the datasets was used as RNAi trigger. This off-target gene silencing risk was confirmed by our experimental analysis with both RNAi transgenic Arabidopsis lines and N. benthamiana plants infected by VIGS constructs. So far, the contribution of silenced off-targets to the silencing phenotype is
not known and needs to be analyzed in exploring target gene function. For example, although the BTII RNAi lines had off-target gene silencing, the transformation recalcitrant phenotype of these lines is not in question as it was confirmed by null mutant lines of the target gene (Hwang and Gelvin, 2004). Nevertheless, no approach used to inactivate gene function is free from potential problems, and the reality of off-target silencing does not override the enormous potential of RNAi as a tool for individual or high throughput studies of gene function. On the other hand, the off-target silencing in PTGS provides an advantage in overcoming gene functional redundancy and in its potential to be applied in heterologous gene silencing across species. Hopefully, a further understanding of the molecular mechanisms of RNAi will add more restraining rules for off-target prediction and reveal possible approaches to overcome off-target effects for target specificity. At this point, our "siRNA Scan" should provide an extremely useful tool in searching the potential off-targets of an RNAi trigger of interest and will also help to design more specific RNAi triggers and appropriate controls in experiments.

MATERIALS AND METHODS

Datasets
Arabidopsis thaliana and rice (Oryza sativa subsp. japonica) genome mRNAs were downloaded from The Institute for Genomic Research (TIGR, http://www.tigr.org/). Gene Indices for 25 plant organisms were downloaded from TIGR. Gene Ontology (GO) database was provided by The Gene Ontology Consortium available at http://geneontology.org. Arabidopsis gene family information and GO annotation were downloaded from The Arabidopsis Information Resource (TAIR) at http://www.arabidopsis.org/. Protein domain data available at http://pfam.wustl.edu/ was also downloaded for off-target function analysis.

Computational off-target estimation
The gene transcript sequence dataset for each organism was searched against itself for contiguous \( \geq 21 \) nt identical or reverse complementary regions using the BLAST algorithm (Altschul et al., 1997). The BLAST output was loaded into a local data warehouse into which Gene Ontology (GO) and gene family databases were also
integrated. Patterns of potential off-target were obtained by querying the database. Possible distinct siRNAs were enumerated from each identical region between a query and an off-target, and efficient siRNAs were also predicted according to the rules by Ui-Tei et al. (Ui-Tei et al., 2004). The rules include 5’ antisense strand (AS) starting with A or U base, 5’ sense strand (SS) starting with G or C, and the first seven bases from 5’ AS end containing at least 3-5 A/U bases. The procedures of the above analyses were modified and implemented in an online tool which is publicly accessible at http://bioinfo2.noble.org/RNAiScan.htm. The input parameters of the tool for analysis of siRNAs and off-targets can be adjusted by users. Additionally, gene function prediction is provided for off-targets by aligning them to Pfam-A seed sequences using BLASTx program. The E-value cutoff is set to $10^{-5}$. When a domain is identified, a local GO database is searched for its available GO annotations in molecular function, biological process and cellular component.

**Plant materials and growth conditions**

*Arabidopsis thaliana* (ecotype Columbia) and *BTII* RNAi transgenic lines (Hwang and Gelvin, 2004) and BTII T-DNA insertion mutant (Salk-032220, *btI*-2, ecotype Columbia) were grown in a growth chamber at 22ºC with 10 hrs daylight. *Nicotiana benthamiana* and the *Agrobacterium tumefaciens* strain GV2260 were used for VIGS analysis. The planting conditions were the same as previously described (Ryu et al., 2004).

**Plasmid construction**

The sequences of TC381 and TC1146 were obtained from the TIGR database of *N. benthamiana*. The cDNA fragments including the sequences with >20 nucleotides continuous identity or complementarity to their potential off-target genes were amplified by reverse transcription (RT)-PCR. The primers for the amplification of specific TC381 and TC1146 fragments are shown in Supplementary Table S2. About 5 µg total RNA from *N. benthamiana* leaves was used for RT at 42ºC with NNPoly(dT)$_{20}$ as the primer for 2 hrs. The RT product was used for PCR amplification for TC381 and TC1146 fragments respectively in a PTC-100™ Peltier Thermal cycler (M. J. Research). The PCR
products were cloned into the pGEM-T-easy vector (Promega, Madison, WI, USA). The clones with inserts that are identical to the area in the target genes were amplified by PCR with primers that contain attB recombination sequences adapted to the previous primer pairs. The PCR products were purified, sequenced and cloned into pTRV2 VIGS vector by using the GATEWAY cloning system with the protocol from the manufacturer (Invitrogen Co., Carlsbad, CA, USA). The recombinant plasmids were named as TRV2::TC381 and TRV2::TC1146.

**Virus-induced gene silencing**
The *Agrobacterium* strain GV2260 containing pTRV1, TRV2::00, TRV2::TC381 and TRV2::TC1146 were grown in an incubator at 28°C on Lubria-bertani (LB) broth with 10 µg L⁻¹ rifampicin and 50 µg L⁻¹ kanamycin for 2 days. Inoculum was prepared with the protocol published previously (Ryu et al., 2004). Leaves of 2 week old *N. benthamiana* plants (2-3 leaf stage) were infiltrated with a 1:1 *Agrobacterium* mixture of either TRV1+TRV2::00 or TRV1+TRV2::TC381 or TRV1+TRV2::TC1146 as described earlier (Liu et al., 2002). The infiltrated plants were grown for 14-18 days for silencing to occur. Leaf samples were then collected for RT-PCR analysis.

**RNA extraction and real time PCR analysis**
For *Arabidopsis*, total RNA was isolated from leaf tissues using TRIZOL Reagent (Invitrogen), followed by RNase-free DNase treatment (Promega). First-strand cDNA was synthesized with the Omniscript RT kit (Qiagen, Crawley, UK) using oligo(dT) 15 according to the manufacturer's instructions. For quantitative PCR, real-time experiments were conducted in an ABI PRISM® 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA), using the intercalation dye SYBRGreen I as a fluorescent reporter. Quantification of PCR products was performed via a calibration curve procedure using *EF-1α* as an endogenous control. The ratio of gene-specific expression to the expression level of the designated calibrator was defined as relative expression using a standard curve method described in User’s Bulletin #2 (Applied Biosystems).
For *N. benthamiana*, total RNA from each plant at 15 days post infiltration (dpi) was extracted using Tri-reagent with the protocol provided by the manufacturer (Molecular Research Center Inc, Cincinnati, OH, USA). First strand cDNA was synthesized with 2 µg of total RNA using primer NNpoly(dT)20 as described above. Semi-quantitative PCR was performed with the modified program from the one described in "Plasmid construction" with annealing temperature at 60ºC and 20 cycles for TC381, 33 cycles for TC1146. The primer sequences are listed in Supplementary Table S3. Further real-time quantitative PCR analysis for each comparison was done with the total RNA of one TRV2::00-infiltrated and one recombinant clone-infiltrated plant, as described above for *Arabidopsis* samples. The analysis was repeated twice with two different groups of plants for each combination. The primer sequences used for real time PCR are listed in Supplementary Table S3.

**Data Analysis**
Data were subjected to analysis of variance using the Student's t-test software of Excel 2003 with 98% confidence.

**ACKNOWLEDGEMENTS**
We thank Drs. Stan Gelvin and Hau-Hsuan Hwang for sending us the *Arabidopsis* seeds of BTI1 RNAi lines and *bti1-2* T-DNA mutant, Dr. S. P. Dinesh-Kumar for providing GATEWAY ready TRV-VIGS vectors, Dr. Choong-Min Ryu for help with infiltrations and discussions, and Drs Elison Blancaflor, Stan Gelvin and Jianzhong Liu for reviewing this manuscript.

**LITERATURE CITED**


Ekengren SK, Liu Y, Schiff M, Dinesh-Kumar SP, Martin GB (2003) Two MAPK cascades, NPR1, and TGA transcription factors play a role in Pto-mediated disease resistance in tomato. The Plant Journal 36: 905-917


Liu Y, Schiff M, Dinesh-Kumar SP (2004) Involvement of MEK1 MAPKK, NTF6 MAPK, WRKY/MYB transcription factors, COII and CTR1 in N-mediated resistance to tobacco mosaic virus. The Plant Journal 38: 800-809


FIGURE LEGENDS

Figure 1. Histograms showing distributions of the numbers of silencing triggers and off-targets in Arabidopsis. (A) number of silencing triggers versus the number of off-targets. (B) number of off-targets versus the number of off-target sites. (C) number of off-target site versus the length of off-target site. Y-axis is in log(10) scale.

Figure 2. Average number of off-target gene family members predicted to be silenced by a trigger from the same family in Arabidopsis. Data for gene families with the same size are combined and averaged.

Figure 3. Comparison of the expression levels of BTII and some of its potential off-targets in wild-type, btii-2 mutant and BTII RNAi transgenic Arabidopsis plants. elf4, initiation factor-4, and Actin II genes were used as controls to show equal loading. Error bars are the stand deviations of three biological replicates of quantitative PCR. *, the expression level of the gene in BTII RNAi transgenic line is significantly lower than that in the wild-type or btii-2 mutant plants with 98% confidence by t-test.

Figure 4. Comparison of phenotypes exhibited by BTII RNAi transgenic lines, btii-2 mutant plants and the wild-type non transgenic plants. All the plants were grown in short day for five weeks and moved to the growth chamber in long day condition for a week. (A), Plants from four independent BTII RNAi transgenic lines. (B), btii-2 mutant plants. (C), Wild-type plants.

Figure 5. Off-target gene silencing in N. benthamiana by VIGS. (A), Comparison of the expression levels of TC381, CK286172, TC10748, CK288691, and β-tubulin in TC381
silenced and non-silenced *N. benthamiana*. (B), Comparison of the expression levels of TC1146, CK282591, CK287535, CK292351, CK296810, TC8666 and $\beta$-tubulin in the *N. benthamiana* plants silenced with TRV2::00 or TRV2::TC1146 construct. Error bars are the standard deviations among three technical replicates of real time PCR. *, the expression level of the gene in TRV2::TC381 or TRV2::TC1146 infiltrated plant is significantly lower than that in TRV2::00 infiltrated plants with 98% confidence by t-test. Both the experiments in (A) and (B) were repeated twice with two individual silenced plants and the results were similar.
Figure 1
Average number of off-target gene family member silenced vs. gene family size (number of family members)
Figure 5
<table>
<thead>
<tr>
<th>Species</th>
<th>Source sequence set</th>
<th>Release Date</th>
<th>Total number of unique gene transcripts</th>
<th>Number of triggers predicted to silence off-targets (% total)</th>
<th>Average number of predicted off-targets per trigger</th>
<th>Average number of predicted off-target regions per off-target</th>
<th>Average length (nt) per off-target region</th>
<th>Average number of efficient siRNAs per trigger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabidopsis thaliana</td>
<td>ATH1</td>
<td>6/10/04</td>
<td>28952</td>
<td>19882 (68.7)</td>
<td>3.9</td>
<td>2.9</td>
<td>41.2</td>
<td>35.2</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>AGI 1</td>
<td>1/29/04</td>
<td>45683</td>
<td>37250 (81.5)</td>
<td>4.3</td>
<td>3</td>
<td>99.2</td>
<td>147.5</td>
</tr>
<tr>
<td>Capsicum annuum</td>
<td>CaGI 1</td>
<td>6/4/04</td>
<td>10712</td>
<td>2788 (26.0)</td>
<td>1.9</td>
<td>2.2</td>
<td>40.7</td>
<td>12.5</td>
</tr>
<tr>
<td>Gossypium spp.</td>
<td>CGI 5</td>
<td>9/16/03</td>
<td>24350</td>
<td>14270 (58.6)</td>
<td>6.9</td>
<td>2.3</td>
<td>74</td>
<td>87.5</td>
</tr>
<tr>
<td>Chlamydomonas reinhardtii</td>
<td>ChrGI 4</td>
<td>1/31/04</td>
<td>30339</td>
<td>16579 (54.6)</td>
<td>6.1</td>
<td>1.9</td>
<td>40.2</td>
<td>12.5</td>
</tr>
<tr>
<td>Glycine max</td>
<td>GmGI 11</td>
<td>9/16/03</td>
<td>67826</td>
<td>49658 (73.2)</td>
<td>9.4</td>
<td>2.1</td>
<td>77.9</td>
<td>108.4</td>
</tr>
<tr>
<td>Helianthus annuus</td>
<td>HaGI 3</td>
<td>9/16/03</td>
<td>20520</td>
<td>11961 (58.3)</td>
<td>15.8</td>
<td>1.4</td>
<td>34.6</td>
<td>38.8</td>
</tr>
<tr>
<td>Hordeum vulgare</td>
<td>HvGI 8</td>
<td>1/30/04</td>
<td>49190</td>
<td>32297 (65.6)</td>
<td>16.3</td>
<td>2.6</td>
<td>99.4</td>
<td>205.9</td>
</tr>
<tr>
<td>Lycopersicon esculentum</td>
<td>LeGI 9</td>
<td>5/19/03</td>
<td>31012</td>
<td>15340 (49.5)</td>
<td>2.4</td>
<td>2.3</td>
<td>60.8</td>
<td>30.9</td>
</tr>
<tr>
<td>Lotus japonicus</td>
<td>LjGI 2</td>
<td>5/16/03</td>
<td>11025</td>
<td>2951 (26.8)</td>
<td>2.6</td>
<td>1.7</td>
<td>35.6</td>
<td>10</td>
</tr>
<tr>
<td>Lactuca sativa</td>
<td>LsGI 2</td>
<td>2/3/04</td>
<td>22185</td>
<td>12345 (55.6)</td>
<td>24.1</td>
<td>1.4</td>
<td>30.4</td>
<td>26.8</td>
</tr>
<tr>
<td>Mesembryanthemum crystallinum</td>
<td>McGI 4</td>
<td>5/15/03</td>
<td>8455</td>
<td>3459 (42)</td>
<td>13.3</td>
<td>2</td>
<td>33</td>
<td>29.4</td>
</tr>
<tr>
<td>Medicago truncatula</td>
<td>MtGI 7</td>
<td>5/19/03</td>
<td>36976</td>
<td>19468 (52.7)</td>
<td>5.6</td>
<td>2.2</td>
<td>49.4</td>
<td>50.6</td>
</tr>
<tr>
<td>Nicotiana benthamiana</td>
<td>NbGI 1</td>
<td>1/28/04</td>
<td>6118</td>
<td>3199 (52.3)</td>
<td>2.8</td>
<td>3.8</td>
<td>53.4</td>
<td>42.1</td>
</tr>
<tr>
<td>Nicotiana tabacum</td>
<td>NtGI 1</td>
<td>6/4/04</td>
<td>10232</td>
<td>2184 (21.3)</td>
<td>1.8</td>
<td>3.1</td>
<td>38.3</td>
<td>15.2</td>
</tr>
<tr>
<td>Allium cepa</td>
<td>OnGI 1</td>
<td>9/17/03</td>
<td>11726</td>
<td>3735 (31.9)</td>
<td>1.7</td>
<td>2.3</td>
<td>54.2</td>
<td>20.2</td>
</tr>
<tr>
<td>Oryza sativa</td>
<td>OsGI 15</td>
<td>5/27/04</td>
<td>88765</td>
<td>45406 (51.2)</td>
<td>22.5</td>
<td>4.4</td>
<td>41.7</td>
<td>174.3</td>
</tr>
<tr>
<td>Pinus spp.</td>
<td>PGI 4</td>
<td>1/28/04</td>
<td>31771</td>
<td>16487 (51.9)</td>
<td>11.7</td>
<td>5.5</td>
<td>38.4</td>
<td>27.5</td>
</tr>
<tr>
<td>Secale cereale</td>
<td>RyeGI 3</td>
<td>1/29/04</td>
<td>5347</td>
<td>1472 (27.5)</td>
<td>2.7</td>
<td>6</td>
<td>32.4</td>
<td>10.7</td>
</tr>
<tr>
<td>Sorghum bicolor</td>
<td>SbGI 8</td>
<td>5/26/04</td>
<td>39148</td>
<td>22039 (56.3)</td>
<td>6.4</td>
<td>1.9</td>
<td>70.3</td>
<td>56.9</td>
</tr>
<tr>
<td>Saccharum officinarum</td>
<td>SoGI 1</td>
<td>1/30/04</td>
<td>95884</td>
<td>55269 (57.6)</td>
<td>24.1</td>
<td>3.5</td>
<td>69.1</td>
<td>205.4</td>
</tr>
<tr>
<td>Solanum tuberosum</td>
<td>StGI 9</td>
<td>5/26/04</td>
<td>32553</td>
<td>20469 (62.9)</td>
<td>4.4</td>
<td>2.9</td>
<td>46.2</td>
<td>42.2</td>
</tr>
<tr>
<td>Triticum aestivum</td>
<td>TaGI 8</td>
<td>1/29/04</td>
<td>123807</td>
<td>89862 (72.6)</td>
<td>19.5</td>
<td>2.7</td>
<td>52.4</td>
<td>102.2</td>
</tr>
<tr>
<td>Theobroma_cacao</td>
<td>TcaGI 1</td>
<td>6/4/04</td>
<td>2539</td>
<td>531 (20.9)</td>
<td>2.6</td>
<td>1.9</td>
<td>37.4</td>
<td>13</td>
</tr>
<tr>
<td>Vitis vinifera</td>
<td>VvGI 3</td>
<td>11/14/03</td>
<td>23109</td>
<td>10548 (45.6)</td>
<td>8.8</td>
<td>1.8</td>
<td>40.3</td>
<td>37.1</td>
</tr>
<tr>
<td>Zea mays</td>
<td>ZmGI 14</td>
<td>1/2/04</td>
<td>56364</td>
<td>40028 (71)</td>
<td>8.9</td>
<td>3</td>
<td>63.1</td>
<td>96.9</td>
</tr>
<tr>
<td>Off-target ID</td>
<td>Identical (I) or reverse complementary (RC) sequence</td>
<td>Blast search</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP167859</td>
<td>I, AGCCTGTTCATAAGGTTCTCGG</td>
<td><em>Arabidopsis thaliana</em> reticulon family protein (RTNLB6) (At3g61560). Contains similarity to DnaJ gene YM8520.10 gb</td>
<td>825566 from <em>S. cerevisiae</em> cosmid gb</td>
<td>Z49705. ESTs gb</td>
<td>Z47720 and gb</td>
<td>Z98799 come from this gene.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP302761</td>
<td>I, AGAAGAAGAAGACTAAGAACG I, AGCCTGTTCATAAGGTTCTCGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC251496</td>
<td>RC, TGATTCTCTTCGTCCTCTTC</td>
<td>GB</td>
<td>AAAM51226 unknown protein *{Arabidopsis thaliana}; similar to UP</td>
<td>Q6Z702 (Q6Z702) Putative 3-isopropylmalate dehydratase large subunit.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC251703</td>
<td>I, CCTGTCATAAGGTTCTCGG; I, TGGTCTATGCCACTATGTTCATT; I, CCAAAGATCTCTGAGTTCATATGCC ATGAAGAACCT; I, TTCTTGACATGGGCAATACAGC</td>
<td>GB</td>
<td>AAP47461.1[32331867]AY164887 RTNLB2 *{Arabidopsis thaliana}; complete. Recently identified as BTI2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC255665</td>
<td>I, TCTTCGTCCTCTTCATCTCTCTTCT</td>
<td>GB</td>
<td>AAD31078.1[4850408]F3F19 Contains PF00097 Zinc finger (C3HC4) ring finger motif. *{Arabidopsis thaliana}, complete.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC258543</td>
<td>I, AGCCTGTTCATAAGGTTCTCGG</td>
<td>GB</td>
<td>AAP47457.1[32331859]AY164883 RTNLB6 *{Arabidopsis thaliana}; partial (90%).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC262843</td>
<td>I, CTCTGTCCTCTCCTCTCATC</td>
<td>UPI</td>
<td>Q49467 (O49467) Hypothetical protein F24J7.50 (Hypothetical protein AT4g19490).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC263798</td>
<td>RC, TTCTTCTCCTCCTCCTATC</td>
<td>GB</td>
<td>AAO50470 unknown protein *{Arabidopsis thaliana}.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC265975</td>
<td>I, AGAAGAAGAAGACTAAGAACG</td>
<td>UPI</td>
<td>Q80799 (O80799) T8F5.5 protein; weakly similar to PIR</td>
<td>S64314 probable membrane protein YGR023w - yeast (<em>Saccharomyces cerevisiae</em>).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC269146</td>
<td>I, CTCTGTCCTCTCCTCTATC</td>
<td>UPI</td>
<td>Q49467 (O49467) Hypothetical protein F24J7.50 (Hypothetical protein AT4g19490).</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC275407</td>
<td>RC, TTCTTCTCCTCCTCCTATC</td>
<td>UPI</td>
<td>Q8VVVC1 (Q8VVVC1) Putative serine threonine kinase.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC275528</td>
<td>I, GAGAAGAAGAAGACTAAGAACG</td>
<td>UPI</td>
<td>QFJR9 (Q9FJR9) Similarity to maturase-related protein, complete.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TC275625</td>
<td>RC, TTCTTCTCCTCCTCCTCATCTCTCT</td>
<td>GB</td>
<td>AAAM51439 unknown protein *{Arabidopsis thaliana}; weakly similar to UP</td>
<td>081812(081812) Auxilin-like protein, partial (13%).</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* TC numbers in bold are the ones used for quantitative RT-PCR.
Table 3. Result from the "siRNA Scan" search for potential off-targets with the cloned sequences of *Nicotiana benthamiana* TC381 and TC1146 (supplementary Figure 1) as trigger sequences

<table>
<thead>
<tr>
<th>Target ID</th>
<th>Annotation of the target</th>
<th>Potential off-target</th>
<th>Identical (I) or reverse complementary (RC) sequence</th>
<th>Annotation of the potential off-target</th>
</tr>
</thead>
<tbody>
<tr>
<td>TC381</td>
<td>similar to TIGR_Osa1</td>
<td>9630. m01295 U2 snRNAP protein A&quot;-Arabidopsis thaliana, partial 79%</td>
<td>CK286172</td>
<td>RC, GGGTGATTCTGGCCCGGGCC TG RC, TATCTTGTAGTAGTTATTAGTATAGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CK288691</td>
<td>RC, TATCTTGTAGTAGTTATTAGTATAGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CK289650</td>
<td>I, TTTCGCGGTCCCCGGGCTTCGTG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC10748</td>
<td>I, TATCTTGTAGTAGTTATTAGTATAGT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC7796</td>
<td>I, TATCTTGTAGTAGTTATTAGTATAGT</td>
</tr>
<tr>
<td>TC1146</td>
<td>similar to UP</td>
<td>Q8H9C6 (Q8H9C6) Pyruvate Decarboxylase (fragment), partial (33%)</td>
<td>CK282591</td>
<td>RC, ATCGCTTTTGCAGACCCGACTAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CK287535</td>
<td>RC, ATCGCTTTTGCAGACCCGACTAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CK292351</td>
<td>RC, TTTGTCTCGGGCCTTACCAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CK296810</td>
<td>RC, ATCGCTTTTGCAGACCCGACTAG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC8666</td>
<td>RC, ATCGCTTTTGCAGACCCGACTAG</td>
</tr>
</tbody>
</table>

TC numbers in bold are the ones used for quantitative RT-PCR. 