THE IMPACT OF siRNA FEATURES ON ITS FIDELITY AND EFFICIENCY IN TARGETING SOLUBLE ACID INVERTASE GENE IN SUGARCANE

SHEREEN K. M. KHALED[1], F. M. ABDEL-TAWAB[2], EMAN M. FAHMY[2], E. M. AHMED[1] AND KH. A. KHALED[3],

1. *Sugar Crop Research Institute (SCRI), Agriculture Research Center*
2. *Genetics Department, Faculty of Agriculture, Ain-Shams University*
3. *Genetics Department, Faculty of Agriculture, Beni-suef University*

Gene silencing can take place through transcriptional or post-transcriptional processes. siRNA are more diverse in plants than in animals, their action can be by degradation of mRNA or activation of epigenetic modifications (Sunkar and Zhu, 2007)

Small interfering RNA (siRNA) and microRNA (miRNA) have emerged as two of the most powerful approaches for crop improvement. RNAi based on siRNA is one of the widely used tools of reverse genetics which aids in revealing gene functions in many species. This technology has been extensively applied to alter gene expression in plants with the aim of achieving desirable traits. As per the FDA guidelines, small RNA (sRNA) based transgenics are much safer for consumption than those over-expressing proteins (Kamthan *et al*., 2015).

Long dsRNA or short-hairpin RNA (shRNA) precursors, which are homologous in sequence to the target gene to be silenced, initiate the process of RNAi. An introduced transgene providing long dsRNA can trigger the RNAi pathway. The dicer enzyme is recruited in the cell to carry out the cleavage of the dsRNA into siRNA that are dsRNAs (21-25 nt). The recruitment of siRNA-induced silencing complex (siRISC) leads to degradation of sense strand of siRNA (having the complementary sequence of the target gene). The siRISC is then incorporated into the antisense strand of siRNA which in association with Argonaute proteins (AGO) and other effector proteins, brings about cleavage of the target mRNA in sequence-specific manner. The activated RISC can repeatedly participate in mRNA degradation and protein synthesis, is inhibited resulting in post transcriptional gene silencing (PTGS) (Hamilton and Baulcombe, 1999; Bernstein *et al*., 2001; Elbashir *et al*., 2001; Tuschl, 2001).

Bart *et al*. (2006) established a system for gene silencing using siRNAs in rice leaf. They concluded that short interfering RNAs (siRNAs) can be used to silence exogenous genes quickly and

Egypt. J. Genet. Cytol., 47:367-375, July, 2018
Web Site (*www.esg.net.eg*)

efficiently. A siRNA targeting luciferase resulted in a significant level of silencing after only three hours and up to 83% decrease in expression.

Activities of sucrose-metabolizing enzymes during the development of internodal storage tissue in sugarcane determine the rate of total sugar accumulation and final sucrose concentration. The hypothesis that sucrose-phosphate synthase and acid invertase play key roles in determining sucrose concentration during maturation in sugarcane internodes has been proved (Lingle, 1999). Disruption of a gene function using reverse genetics is a very useful and relatively simple way to analyze gene functions and achieve crop trait improvements (Kumar *et al*., 2012). Different approaches are used for gene silencing either at transcriptional or post transcriptional level or both. Antisense, VIGS, hpRNAi and amiRNA are some of the most commonly used small RNA mediating gene silencing pathways in plants. All silencing pathways have the route with 21-24 nt sequence which conduct the silencing action transcriptionally or post-transcriptionally. Recently, artificial siRNA has been used in plants to silence the gene directly and easily.

Invertases are some of the key regulatory enzymes involved in various physiological and biochemical reactions in sugarcane. Different developmental stages like early vegetative growth and stem development, grand growth period, ripening and transition of the cane from source to sink form are mediated by an intrinsic regulation of invertase activity. The regulation of invertases *in vivo* is not fully understood, though there are several mechanisms including post translational suppression by inhibitory proteins. These inhibitory proteins are not yet characterized in sugarcane (Prathima *et al.*, 2014).

The overall aim of this study was to design an effective siRNA sequence that could down-regulate the expression of soluble acid invertase enzyme in some Egyptian sugarcane varieties.

MATERIALS AND METHODS

1- Short hairpin RNA (shRNA) designing

The complete sequence of soluble acid invertase gene (SAI) was obtained from NCBI database with accession numbers JQ982495.1, JQ406877.1, JQ406876.1 and JQ406875.1. For the design of siRNA molecule, the sequence of SAI1 gene with accession number (JQ406875.1) was used. The design requires four separate steps.

Step 1: siRNA design

There are different programs and websites is one word to detect the siRNA sequences in a specific gene but most of them are specialized for rats and humans. PSSRNAit is a website specialized for plants and can detect the expected off-targeting by searching through different cDNA/transcript libraries for sugarcane, *Saccharum officinarum*. Unigene, DFCI

Gene Index (SOGI),"version3, released on 2010_04_09" was the used library. So, this web site (http://plantgrn.noble.org/pssRNAit/Home.gy) was used to detect the siRNA sequences for the *soluble acid invertase* gene in sugarcane using the sequence of *SAI1* gene with accession number (JQ406875.1).

Step (2): NCBI blast of siRNA

Although the PSSRNAit website represents the expected numbers of off-targets, a BLASTn is still necessary for each siRNA sequence. The blasts on NCBI database minimize the possibility of damaging any other gene.

Step (3): Secondary structure prediction

The designed siRNA molecules should be tested for the secondary structure formation to increase their fidelity and decrease the possibilities of non-functional molecules inside the cell. There are many mfold programs which can predict the secondary structure. The Oligonucleotide Properties Calculator, Oligo Calculator version 3.26, was used http://www. basic.northwestern.edu/ biotools/oligocalc.html.

Step (4): ShRNA design

There are different websites used to convert siRNA to shRNA molecule by choosing the suitable loop sequence. The Biosettia.com website was used.

2- Testing the efficacy of the designed siRNA

In-planta transformation experiment was conducted using the siRNA1 sequence which was ligated in the expression vector pFGC5941. The sugarcane variety G99/103 was used in this experiment. The sugarcane variety was transformed according to the modified method of Mayavan *et al*. (2015). The details of the experiment were reported by Khaled *et al.* (2018). To test the fidelity of siRNA sequence, RT-PCR was used to estimate the down regulation in the expression of the targeted gene.

RESULTS AND DISCUSSION

1- Short hairpin RNA (shRNA) designing

The design of shRNA from *soluble acid invertase (SAI1)* gene with accession number (JQ406875.1) requires four steps as mentioned previously. The results of each step are shown as following;

Step 1: siRNA design

By using PSSRNAit web site, 234 candidate siRNA sequences were obtained as shown in Table (1). The first four siRNA sequences were tested carefully before one of them was used to confirm their ability to silence the targeted gene without affecting any other gene. There are four main parameters that should be taken into consideration to evaluate siRNA sequences; efficiency, RNA-induced silencing complex (RISC) binding antisense score, target accessibility and the number of off-targets. Despite the expectation of the antisense meeting in 2001 that there is no need to select for optimal siRNA, it was proven later that

the silencing efficiency varied widely against the same target (Holen *et al.*, 2002). It was noted that the efficiency values ranged from zero to ten and the greater is better (Table 1). So, the first four siRNA molecules efficiencies which ranged from 9.75 to 10.31 can be described as desirable for efficiency value sequences.

A significant progress towards the design of active siRNA could be achieved when the incorporation of the two strands of siRNA were taken into consideration. It was clear that functional duplexes displayed a lower relative thermodynamic stability at the $5^{/}$ end of the antisense strand than the nonfunctional duplexes according to the hypothesis that strand incorporation into RISC is determined by an RNA helicase that initiates dissociation of the siRNA duplex at the end with the lower thermodynamic stability (Schwarz *et al.*, 2003). So, the RISC binding antisense score is an important parameter to increase the activity of silencing. Table (1) showed that the second siRNA has the highest score and satisfactory ability of binding to the RISC complex.

The thermodynamic properties of the siRNA as well as the structure of target RNA both influenced the efficiency of siRNA. So, siRNA with favorable thermodynamic properties may be incapable of inhibiting gene expression due to tight secondary structure of the target. Therefore, the target accessibility may further modulate the silencing efficiency (Luo and Chang, 2004). These observations were not in agreement with the results of Pascut *et al.* (2015) who found that there was no association between the number of unpaired bases in the mRNA target region and siRNA efficiency. Also, they showed that there was no association between the number of consecutive unpaired nucleotides per targeted local mRNA secondary structure and consequently, no association between silencing efficacy and targeted mRNA local structure. But their results detected that RBPs (Ribosomal binding proteins), which can mask siRNA target site affect mainly the siRNA efficiency. Despite this, the target accessibility (TA) is a measurable value in the siRNA design. It was noticed that, the highest value of TA was for the fourth sequence, 22.256, and the lowest value was for the second one, 17.913. Therefore, all the first four sequences have unsatisfactory values where the values should be ranged from zero to 25, the lower being the better.

Herein, the number of off-targets varied for the first four siRNAs sequences (Table 1). The lowest value is better as it minimizes the probability of missing its hitting point. Whatever were the thermodynamic properties or the efficiency of siRNA sequence, it is worth nothing without appropriate off-targeting criteria. As silencing of one stone of the system is a crucial step, most of siRNA designing websites carry out a BLAST test in different databases. Despite this, separate BLASTn's were carried out in the next step.

Step (2): NCBI blast of siRNA

To minimize the off-targeting, a BLASTn was carried out for the first four siRNA sequences. It was noticed that all the four siRNA scored a high degree of sequence homology (100% iden.) with *Saccharum* hybrid cultivar GT28 *soluble acid invertase* isoforms and even with other cultivars like Pindar and FN-28. *Sorghum* is the diploid model used for sugarcane studying due to their close synteny (Guimaraës *et al*., 1997; Wang *et al*., 2010). So, the homology with sorghum invertases is expected as the homology between its genome and sugarcane reached up to 95%. All the four siRNA sequences showed the same E value = 0.027 and max score = 42.1 which are acceptable values.

Step (3): Secondary structure prediction

As the symmetric strand incorporation into RISC is controlled by the target mRNA, the thermodynamic properties of the siRNA also have a major effect. So, the mfold analysis has been carried for the aforementioned siRNA sequences. There are many features that should be emphasized in these results. The number of formed hairpin structure for each siRNA and their ΔG and ΔH values are the first two major parameters to be scoped. The ΔG variation of breaking target structures and of intramolecular secondary structural formation, including self-structure formation in the target and in siRNA is an accounted parameter. It was found that the silencing efficacy was positively associated with the stability of the siRNA-mRNA duplex and negatively associated with the stability of the siRNA and mRNA secondary structures (Pascut *et al.*, 2015).

As it is evident from Table (2) the siRNA1 and siRNA4 possess two hairpin structures. On the other hand, siRNA2 is assumed to produce only one structure and all ΔG and ΔH values for the three siRNA are satisfactory. The siRNA3 sequence was eliminated as it was expected to initiate eight structures which is a high number even though their ΔG values are low.

The third feature that should be taken into consideration was the accessible end of the secondary structure of siRNA molecules. It was proven that guide strands that possess freely available terminal nucleotides at the 3’ end increase the efficiency of silencing (Patzel *et al.*, 2005). As shown in the results of the mfold analysis, all the predicted structures possess free ends except two forms of siRNA3 which have a closed end and were excluded.

Step (4): shRNA design

According to the results of the aforementioned analysis, the first siRNA which was expected to target bases from 2161 to 2181 [siRNA1: 5’UUGUUGAAGAGGAACACGCCG3’ with efficiency = 10.31] was used. The loop sequence (TTGGATCCAA) was used in this study.

2- Evaluation siRNA efficacy

The transgenic sugarcane plants obtained from the *in-planta* transformation experiment were tested using RT-PCR by Khaled *et al.* (2018). The changes in *SAI* expression levels were mostly around ten folds decrease in all transgenic plants with minor differences between them (Fig. 1). These results confirm the fidelity of siRNA1 and the other two promising siRNA sequences can be evaluated further to have the highest silencing percentage for the targeting gene.

Finally, Crop improvement through biotechnological approaches has been accelerated widely in recent years. Using functional genomics and bioinformatic tools are promising prospects in this era. Using siRNA as a tool for silencing in sugarcane can be considered a promising tool. So, the applicability of using siRNA approaches proved to be satisfactory. The other two siRNA sequences designed in our experiment should be tested to obtain more transgenic clones with low activity of invertase. This could enhance the production of sucrose in sugarcane plants to help in decreasing the gap between the production and consumption, and consequently could save the hard currency used in importing table sugar

SUMMARY

Bioinformatic tools to design a suitable siRNA are an essential step before conducting any laboratory experiment. The identification of highly active siRNA is the corner stone in any silencing application. Five separate steps were used in these studies to test the siRNA efficacy *in-silico* to evaluate the validation of the siRNA sequence targeting soluble acid invertase gene in sugarcane. The used programs and websites for designing siRNA molecule to silence SAI gene in sugarcane confirmed their validity in designing an efficient, less laboring and accurate siRNA sequence. The subsequent five steps were essential to avoid any side effect on the overall physiological pattern in the plant.

REFERENCES

Bart, Rebecca, M. Chern, C. Park, Laura Bartley and Pamela C. Ronald (2006). A novel system for gene silencing using siRNAs in rice leaf and stem-derived protoplasts. Plant Methods, 2: 13-17.

Bernstein, E., A. A. Caudy, S. M. Hammond and G. J. Hannon (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature, 409: 363-366.

Elbashir, S. M., W. Lendeckel and T. Tuschl (2001). RNA interference is mediated by 21- and 22- nucleotide RNAs. Genes Dev., 15: 188-200.

Guimaraès, Claudia T., G. R. Sills and B. W. S. Sobral (1997). Comparative mapping of Andropogoneae: *Saccharum* L. (sugarcane) and its relation to sorghum and maize.

Proc. Natl. Acad. Sci. USA, 94: 14261-14266.

Hamilton, A. J. and D. C. Baulcombe (1999). A novel species of small anti-sense RNA in posttranscriptional gene silencing. Science, 286: 950-952.

Holen, T., M. Amarzguioui, M. T. Wiiger, E. Babaie and H. Prydz (2002). Positional effects of short interfering RNAs targeting the human coagulation trigger tissue factor. Nucleic Acids Research, 30: 1757-1766.

Kamthan, A., A. Chaudhuri, M. Kamthan and A. Datta (2015). Small RNAs in plants: recent development and application for crop improvement. Front. Plant Sci., 6: 208-212.

Khaled, Shereen K. M., F. M. Abdel-Tawab, Eman M. Fahmy, E. A. M. Amer and K. A. Khaled (2018). The siRNA efficacy of soluble acid invertase down-regulation in Sugarcane (*Saccharum* spp.). Arab Univ. J. Agri. Sci., Special issue, 26(2).

Kumar, P., M. Kamle and A. Pandey (2012). RNAi: new era of functional genomics for crop improvement. In: Goyal A, Maheshwari P(ed) Frontiers on recent developments in plant science, Bentham Sci., 1: 24-38.

Lingle, Sarah E. (1999). Sugar metabolism during growth and development in sugarcane internodes. Crop Sci., 39: 480-486.

Luo, K. Q. and D. C. Chang (2004). The gene-silencing efficiency of siRNA is strongly dependent on the local structure of mRNA at the targeted region. Biochemical and Biophysical Research Communications, 318: 303-310.

Mayavan, S., K. Subramanyam, B. Jaganath, D. Sathish, M. Manickavasagam and A. Ganapathi (2015). *Agrobacterium*-mediated *in-planta* genetic transformation of sugarcane setts. Plant Cell Rep., 34: 1835-1848.

Patzel, V., S. Rutz, I. Dietrich, C. Koberle, A. Scheffold and She Kaufmann (2005). Design of siRNA producing unstructured guide RNAs results in improved RNA interference efficiency. Nature Biotechnology, 23: 1440-1444.

Prathima, P. T., T. V. Suparna, S. Anishma, R. Punnya and K. Ramalakshmi (2014). Cloning and characterization of a differentially regulated *invertase inhibitor* gene during sucrose accumulation in Sugarcane. Journal of Sugarcane Research, 4: 21-28.

Pascut, D., G. Bedogni and C. Tiribelli (2015). Silencing efficacy predic-

tion: a retrospective study on target mRNA features. Bioscience Reports, 35-42.

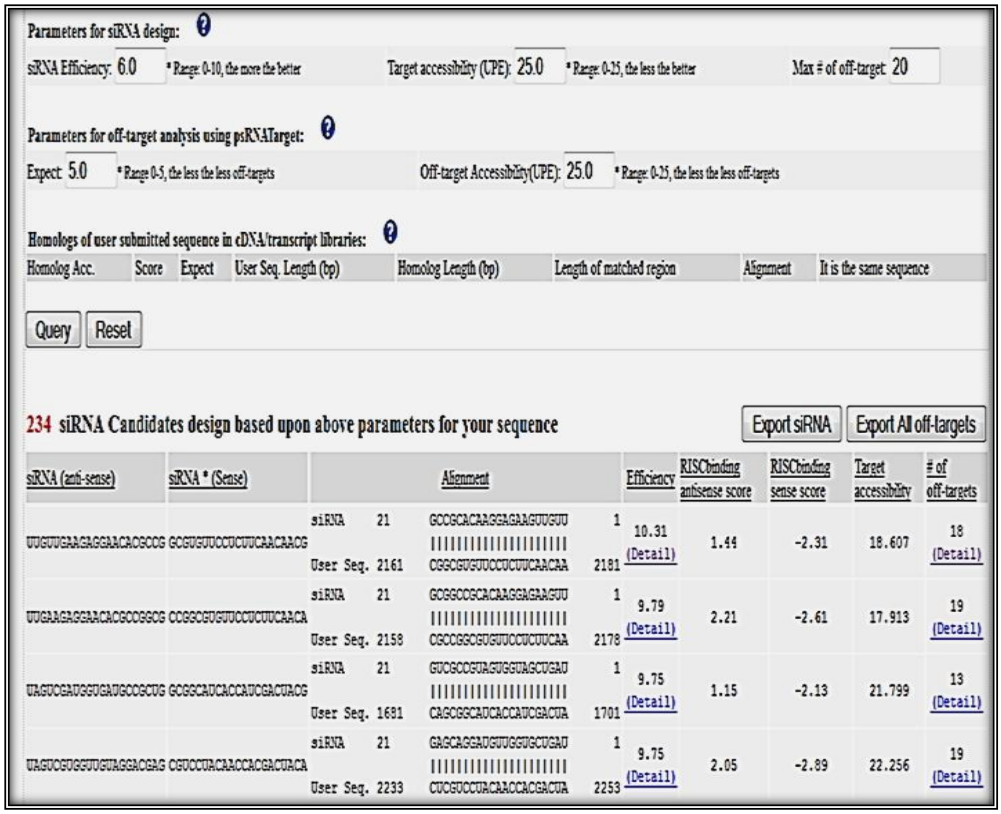
Schwarz, D. S., G. Hutvágner, T. Du, Z. Xu, N. Aronin and P. D. Zamore (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell, 115: 199-208.

Sunkar, R. and J. K. Zhu (2007). Micro RNAs and Short-interferring RNAs in plants. J. Integrative Plant Biology, 49: 817-826.

Tuschl, T. (2001). RNA interference and small interfering RNAs. Chem. Biochem., 2: 239-145.

Wang, J., B. Roe, S. Macmil, Q. Yu, J. E. Murray, H. Tang, C. Chen, F. Najar, G. Wiley, J. Bowers, Marie-Anne Van Sluys, D. S. Rokhsar, M. E. Hudson, S. P. Moose, A. H. Paterson and R. Ming (2010). Micro-collinearity between autopolyploid sugarcane and diploid sorghum genomes. BMC Genomics, 11: 261-278.

Table (1): Candidate siRNA sequences resulting from analysis of *Saccharum* hybrid cultivar GT28 soluble acid invertase1 (SAI1) mRNA.



Parameters for siRNA design:

siRNA Efficiency: 6.0 * Range: 0-10, the more the better | Target accessibility (UPE): 25.0 * Range: 0-25, the less the better | Max # of off-target: 20

Parameters for off-target analysis using psRNATarget:

Expect: 5.0 * Range 0-5, the less the less off-targets | Off-target Accessibility(UPE): 25.0 * Range: 0-25, the less the less off-targets

Homologs of user submitted sequence in cDNA/transcript libraries:

Homolog Acc.	Score	Expect	User Seq. Length (bp)	Homolog Length (bp)	Length of matched region	Alignment	It is the same sequence

Query | Reset

234 siRNA Candidates design based upon above parameters for your sequence — Export siRNA | Export All off-targets

siRNA (anti-sense)	siRNA * (Sense)	Alignment				Efficiency	RISCbinding antisense score	RISCbinding sense score	Target accessibility	# of off-targets
UUGUUGAAGAGGAACACGCCG	GCGUGUUCCUCUUCAACAACG	siRNA	21	GCGGCACAAGGAGAAGUUGUU	1	10.31 (Detail)	1.44	-2.31	18.607	18 (Detail)
		User Seq.	2161	CGGCGUGUUCCUCUUCAACAA	2181					
UUGAAGAGGAACACGCCGGCG	CGGCGUGUUCCUCUUCAACA	siRNA	21	GCGGCCGCACAAGGAGAAGUU	1	9.79 (Detail)	2.21	-2.61	17.913	19 (Detail)
		User Seq.	2158	CGCCGGCGUGUUCCUCUUCAA	2178					
UAGUCGAUGGUGAUGCCGCUG	GCGGCAUCACCAUCGACUACG	siRNA	21	GUCGCCGUAGUGGUAGCUGAU	1	9.75 (Detail)	1.15	-2.13	21.799	13 (Detail)
		User Seq.	1681	CAGCGGCAUCACCAUCGACUA	1701					
UAGUCGUGGUUGUAGGACGAG	CGUCCUACAACCACGACUACA	siRNA	21	GAGCAGGAUGUUGGUGCUGAU	1	9.75 (Detail)	2.05	-2.89	22.256	19 (Detail)
		User Seq.	2233	CUCGUCCUACAACCACGACUA	2253					

Table (2): The number of formed secondary structures for siRNA sequences and their ΔG and ΔH values.

siRNA 5' ⟶ 3'	No. of hairpin structures	ΔG and ΔH value
SiRNA1: UUGUUGAAGAGGAACACGCCG	Two	-2.10 and -28.00 -1.70 and -37.70
SiRNA2: UUGAAGAGGAACACGCCGGCG	One	-2.10 and -28.00
SiRNA3: UAGUCGAUGGUGAUGCCGCUG	Eight	Discarded
SiRNA4: UAGUCGUGGUUGUAGGACGAG	Two	-3.40 and -46.50 -2.40 and -43.10

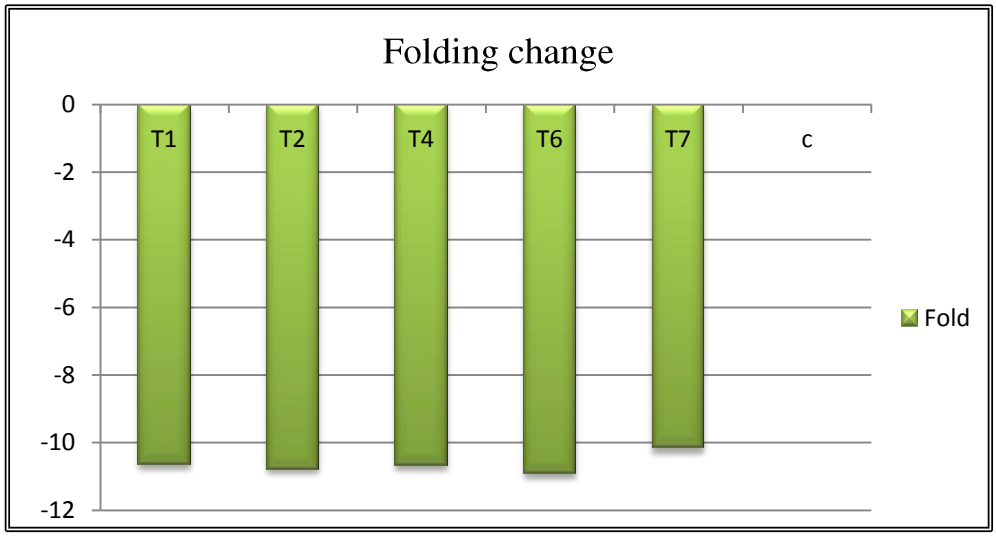


Fig. (1): The fold change in gene expression level of *SAI* gene for silenced sugarcane transgenic plants using siRNA approch compared to control.