# A Novel Computational Approach to the Silencing of Sugarcane Bacilliform Guadeloupe A Virus Determines Potential Host-Derived MicroRNAs in Sugarcane (*Saccharum officinarum* L.)

**Fakiha Ashraf** [Equal first author, 1], **Muhammad Aleem Ashraf** [Corresp., Equal first author, 2, 3], **Xiaowen Hu** [4], **Shuzhen Zhang** [Corresp. 1]

[1] Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou, Hainan, China

[2] Haikou Experimental Station, Chinese Academy of Tropical Agricultural Sciences, Haikou, Hainan, China

[3] Department of Plant Breeding and Genetics, University College of Agriculture and Environmental Sciences, The Islamia University of Bahawalpur, Baghdad-Ul-Jadeed Campus, Bahawalpur, South Punjab, Pakistan

[4] Zhanjiang Experimental Station, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang, Guandong, China

Corresponding Authors: Muhammad Aleem Ashraf, Shuzhen Zhang
Email address: muhammad.aleemashraf@iub.edu.pk, zhangshuzhen@itbb.org.cn

Sugarcane Bacilliform Guadeloupe A Virus (SCBGAV, genus *Badnavirus,* family Caulimoviridae) is an emerging, deleterious pathogen of sugarcane which presents a substantial barrier to producing high sugarcane earnings. The circular, double-stranded (ds) DNA genome of SCBGAV (7.4 Kb) is composed of three open reading frames (ORF) that replicate by a reverse transcriptase. In the current study, we used miRNA target prediction algorithms to identify and comprehensively analyze the genome-wide sugarcane (*Saccharum officinarum* L.)-encoded microRNA (miRNA) targets against the SCBGAV. A total of 28 potential mature target miRNAs were retrieved from the miRBase database and were further analyzed for hybridization to the SCBGAV genome. Multiple computational approaches—including miRNA-target seed pairing, multiple target positions, minimum free energy, target site accessibility, maximum complementarity, pattern recognition and minimum folding energy for attachments— were considered by all algorithms. Only 4 sugarcane miRNAs are selected for SCBGAV silencing. Among those 4, sof-miR396 was identified as the top effective candidate, capable of targeting the vital ORF3 which encodes polyprotein of the SCBGAV genome. miRanda, RNA22 and RNAhybrid algorithms predicted hybridization of sof-miR396 at common locus position 3394. A Circos plot was created to study the network visualization of sugarcane-encoded miRNAs with SCBGAV genes determines detailed evidence for any ideal targets of SCBGAV ORFs by precise miRNAs. The present study concludes a comprehensive report towards the creation of SCBGAV-resistant sugarcane through the expression analysis of the identified miRNAs.

# A Novel Computational Approach to the Silencing of Sugarcane Bacilliform Guadeloupe A Virus Determines Potential Host-Derived MicroRNAs in Sugarcane (*Saccharum officinarum* L.)

Fakiha Ashraf[1, †], Muhammad Aleem Ashraf[2, 3 †,*], Xiaowen Hu[4] and Shuzhen Zhang[1,*]

[1]Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China

[2]Haikou Experimental Station, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China

[3]Department of Plant Breeding and Genetics, University College of Agriculture and Environmental Sciences, Islamia University of Bahawalpur, Baghdad-Ul-Jadeed Campus, Bahawalpur 63100, Pakistan

[4]Zhanjiang Experimental Station, Chinese Academy of Tropical Agricultural Sciences, Zhanjiang, 524091, China

[†] These authors contributed equally to this work

[*] Correspondence: muhammad.aleemashraf@iub.edu.pk , zhangshuzhen@itbb.org.cn

## Abstract

Sugarcane Bacilliform Guadeloupe A Virus (SCBGAV, genus *Badnavirus,* family Caulimoviridae) is an emerging, deleterious pathogen of sugarcane which presents a substantial barrier to producing high sugarcane earnings. The circular, double-stranded (ds) DNA genome of SCBGAV (7.4 Kb) is composed of three open reading frames (ORF) that replicate by a reverse transcriptase. In the current study, we used miRNA target prediction algorithms to identify and comprehensively analyze the genome-wide sugarcane (*Saccharum officinarum* L.)-encoded microRNA (miRNA) targets against the SCBGAV. A total of 28 potential mature target miRNAs were retrieved from the miRBase database and were further analyzed for hybridization to the SCBGAV genome. Multiple computational approaches—including miRNA-target seed pairing, multiple target positions, minimum free energy, target site accessibility, maximum complementarity, pattern recognition and minimum folding energy for attachments— were considered by all algorithms. Only 4 sugarcane miRNAs are selected for SCBGAV silencing. Among those 4, sof-miR396 was identified as the top effective candidate, capable of targeting the vital ORF3 which encodes polyprotein of the SCBGAV genome. miRanda, RNA22 and RNAhybrid algorithms predicted hybridization of sof-miR396 at common locus position 3394. A Circos plot was created to study the network visualization of sugarcane-encoded miRNAs with SCBGAV genes determines detailed evidence for any ideal targets of SCBGAV ORFs by precise miRNAs. The present study concludes a comprehensive report towards the creation of SCBGAV-resistant sugarcane through the expression analysis of the identified miRNAs.

## Introduction

Sugarcane bacilliform virus (SCBV, genus: *Badnavirus*, family Caulimoviridae) is a circular, non-enveloped bacilliform, monopartite DNA virus that harbors about 7.5 kb double-stranded (ds) DNA with three open reading frames (ORFs) (Sun et al. 2016). SCBV was first reported in the sugarcane cultivar (B34104) in South American continent (Cuba, 1985) and disseminated into many sugarcane (*Saccharum officinarum* L.)-growing regions in the globe (AUTREY et al. 1992). Sugarcane Bacilliform Guadeloupe A Virus (SCBGAV) was first identified as a sugarcane R570 hybrid variety in Guadeloupe in 2011 (Muller et al. 2011). SCBGAV was classified as a new species in the genus badnavirus under the SCBV-A group as assigned by ICTV(Adams et al. 2016). The disease symptoms appear in the form of chlorosis, leaf freckle and mottling. Some infected plants were also observed to exhibit no symptoms. Recently in China, SCBV-infection resulted in reduced sucrose content, juice, stalk weigh, purity and gravity in sugarcane plants(Ahmad et al. 2019).

SCBGAV is a mealybug-transmitted monopartite badnavirus that infects sugarcane and is composed of a 7444 bp circular, dsDNA genome containing three ORFs on the positive strand that replicate by a virus-encoded reverse transcriptase (RT). The two small proteins of sizes (176-185 and 122-135 amino acids), were encoded by the ORF1 and ORF2 respectively. The ORF3 was a key component of SCBGAV genome, encoded by a polyprotein (1786-1933 amino acids) (Muller et al. 2011). Sugarcane (*Saccharum officinarum* L.) has inherited an active immunity that is composed of microRNAs (miRNAs) to combat infection. The miRNAs are a class of 21-23 nucleotide-long, endogenous, noncoding RNAs that govern gene regulation and expression at the post-transcriptional level. In plants, this mechanism occurs through the process of mRNA cleavage or translational repression. They are synthesized after the processing of hairpin miRNA precursors using an RNase-III-like enzyme (Dicer) to control gene expression (Brodersen & Voinnet 2006).

The development of artificial microRNAs (amiRNAs)-mediated gene silencing method offers a highly precise, targeted approach that gives several advantages over RNAi-mediated gene silencing strategy. Design, construction and validation of the amiRNAs depend upon the assembly of an endogenous precursor miRNA, which is substituted with a selected miRNA nucleotide transcript complementary to the target sequence(Niu et al. 2006; Ossowski et al. 2008).

The objective of the present research work is to predict sugarcane-encoded miRNAs that have the potential to develop resistance against badnaviruses, especially the SCBGAV. The major research objective was designed to identify the potential host-derived miRNAs in the SCBGAV genome and to screen the most promising miRNAs to understand the complex host-virus interactions. The novel amiRNA silencing method has been adopted for the first time to explore as a source of creating resistance to a monopartite *badnavirus*. The predicted miRNAs may provide help for the generation of SCBGAV-resistant sugarcane plants through genetic engineering in the future.

## Materials & Methods

### Retrieval of Biological Data

The sugarcane (*Saccharum-officinarum* L) and (*Saccharum spp*.) plant miRNA sequences were accessed from the microRNA database, miRBase (http://www.mirbase.org/cgi-bin/browse.pl/) (Griffiths-Jones et al. 2006). A total of 28 miRNAs were retrieved. The nucleotide sequence of SCBGAV dsDNA genome (accession number NC_038382.1) composed of 7444bp and was retrieved from the NCBI database. Four computational algorithms — named miRanda, RNA22, RNAhybrid and psRNATarget — were selected for the screening of mature miRNAs of sugarcane against the SCBGAV genome to identify the miRNAs target positions (Table1). Nucleotide sequences of the SCBGAV genome and sugarcane miRNAs were recorded in FASTA format to process further using computational algorithms with defined parameters.

### miRanda

miRanda is a computational algorithm used to predict and identify the potential plant genomic miRNA targets (John et al. 2004). miRanda software was downloaded and run after defining standard settings. These include: Gap Open Penalty = −9.0, Score Threshold = 130, Minimum free energy (MFE) threshold (= −15.00 kcal/mol). MFE is an important statistical parameter to screen potential targets.

### RNA22

RNA22 is a novel diverse web server (http://cm.jefferson.edu/rna22v1.0/), designed to implement a pattern-based approach to detect potential miRNA target sites. The minimum folding energy (MFE) is a key parameter to screen possible miRNA without a cross-species conservation filter (Loher & Rigoutsos 2012). Specificity and sensitivity values were selected at 61% and 63%, respectively. The value of the seed size was selected as 7 with unpaired base 1 permitted inside seed region. There was no limit set to the maximum number of G:U wobbles in the seed region. The paired bases with minimum number and MFE were selected at 12 and -13.50 kcal/mol, respectively.

### RNAhybrid

RNAhybrid is a novel algorithm used to predict hybridization of miRNA and mRNA that is based on MFE (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid). Site complementarity and MFE are the unique parameters of RNAhybrid algorithm. MFE was set at threshold of −20kcl/mol. The other filters remained fixed as default parameters(Krüger & Rehmsmeier 2006).

### psRNATarget

psRNATarget is a new web server (http://plantgrn.noble.org/psRNATarget/) that was developed to predict small RNA (sRNA) in plants. It was used to analyze complementary matching between target mRNA sequence and sRNA sequence based on a scoring schema (Dai et al. 2018). Evaluation of target-site accessibility on the basis of UPE (unpaired energy) is another important feature of the psRNATarget algorithm. The standard parameter set for our analysis was as follows: penalty for (extending gap = 0.5, opening gap = 2, G.U pair = 1, other mismatches= 1), HSP size=19, seed region= 2-7 nucleotides. The minimum expectation score was 7.0.

**Development of a Circos Map between miRNA and Target**

A Circos plot was developed between sugarcane-encoded miRNA and SCBGAV genes by applying the Circos algorithm (Krzywinski et al. 2009).

**Statistical Analysis**

Predicted miRNA data obtained by applying four diverse algorithms was further analyzed using R-Language (v3.1.1, software version 3.5.1) (Gandrud, 2013). The graphical representation of the predicted miRNAs was processed using in house-scripts (readxl and ggplot2 packages).

## Results

**Prediction of Sugarcane-encoded miRNA Targets in the SCBGAV ORF1**

There is little information available about the function of this protein. ORF1 (538-1071 nucleotides) encodes a hypothetical protein represented as P1 (177 amino acids). RNAhybrid predicted binding of sof-miR159e at locus 733, ssp-miR169 at locus 752, ssp-miR437 (a, c) at locus 978, and ssp-miR444 (a, b, c-3p) at locus 774 (Fig.1c). In addition, psRNATarget predicted hybridization of ssp-miR166 at locus 981, ssp-miR444a at locus 775, and ssp-miR444b at position 756 (Fig.1d)

**Identification of miRNA targets in the SCBGAV ORF2**

The DNA binding protein, represented as P2, is encoded by the ORF2. ORF2 has the least number of predicted targets by sugarcane miRNAs; only three miRNA of sugarcane (sof-miR167 (a, b) and ssp-miR528) were targeted at common locus 1301, indicated by RNA22 (Fig.1b). Similarly, RNAhybrid predicted binding of sof-miR167 (a, b) at locus 1304 and ssp-miR1432 at locus 1301 (Fig.1c).

**miRNA Target prediction in the ORF3**

The polyprotein is encoded by the ORF3 which contains several functional units. These proteins are coat, movement, aspartic protease, RT, and ribonuclease H. For the ORF3 of SCBGAV, seven different kinds of miRNAs were predicted by miRanda: (sof-miR159 (a, b, c, d, and e), sof-miR168a, sof-miR396, ssp-miR166, ssp-miR827, ssp-miR1128 and ssp-miR1432) (Fig.1a). RNA22 predicted hybridization of the following miRNAs: (sof-miR168 (a, b), sof-miR396, sof-miR408 (a, b, c and d) and ssp-miR444 (a, b) (Fig.1b). Suitable miRNAs for targeting ORF3 were hypothesized by RNAhybrid to interact with the ORF3 of SCBGAV. These are sof-miR159 (a, b, d), sof-miR168a, sof-miR396, sof-miR408 (a, b, c, d and e), ssp-miR166, ssp-miR827, and ssp-miR1128 (Fig.1c). Moreover, sof-miR159 (a, b, c, d), sof-miR167 (a, b), sof-miR396, ssp-miR528, ssp-miR444b and ssp-miR444c-3p were identified by psRNATarget (Fig.1d).

**Visualization of miRNA Target Interaction**

A Circos plot was used to study the visualization of sugarcane miRNAs with SCBGAV-targets (ORFs) that display particular evidence for potential targets. We have for the first time reported sugarcane-encoded miRNAs and their targets simultaneously constructed in this manner (Fig. 2).

**Sugarcane miRNAs (Predicted by a Consensus of Algorithms) for SCBGAV Silencing**

Among all the predicted miRNAs for SCBGAV silencing, only four miRNAs (sof-miR167 (a, b), sof-miR396 and ssp-miR528) were predicted by all of the algorithms used (Fig.3 and Table 2). Moreover, six miRNAs (sof-miR167 (a, b) at locus 5846 and 1310, sof-miR168a at locus

5506, sof-miR396 at locus 3394, ssp-miR444a at locus 775 and ssp-miR1128 at locus 6148) were predicted at the common locus by at least two of the algorithms used (Fig.4). Out of 28 sugarcane miRNAs, only one miRNA (sof-miR396) was predicted by at least three algorithms used to have binding site at the same locus position 3394 (Table3).

## Discussion

For possible miRNA target prediction in the genome of SCBGAV, a combination of the aforementioned computational tools was used in order to filter out the false positive results and to increase the accuracy of miRNA target prediction. miRanda was implemented to validate various parameters, from target site conservation to whole genome prediction of miRNA target genes. Then, RNAhybrid and psRNATarget were used, both of which are strongly prescribed for plant miRNA target identification. RNA22 is a novel algorithm that applies an *in-silico* strategy that is highly divergent in comparison to other algorithms. Pattern-based recognition is the key feature of this algorithm.

The results from this study suggest that SCBGAV is susceptible to targeting by consensus sugarcane-encoded miRNAs. The genome components of SCBGAV (ORF1, ORF2 and ORF3) seemed to be principally prone sites for sugarcane-miRNA regulation. While *in vivo* demonstration requires validating functional efficacy, the degree of complementarity between target mRNA and miRNA concludes the fate of the predicted sites. A full complementarity binding between target mRNA and miRNA sequence results in endonucleolytic cleavage and disruption. Contrary to this, partial target-site complementarity characteristically down-regulates target-gene expression by suppressing translation of target mRNA (Liu et al. 2017).

The present study identifies suitable sugarcane-encoded miRNAs to exhibit a stronger degree of target-site complementarities within the ORF1, ORF2 and ORF3 of SCBGAV. These predicted miRNAs may be utilized to develop effective amiRNA constructs, which could be used to enhance the sugarcane immunity to SCBGAV. Pairing multiple miRNAs with a single mRNA induces effective RNA silencing (Doench & Sharp 2004). In the current study, we have predicted several miRNA targets which were associated with SCBGAV (ORF1, ORF2, and ORF3) genes at multiple loci. A deeper understanding of these vital ORFs involved in SCBGAV epidemic via miRNA-mediated control of gene expression would significantly assist in the development of molecular approaches to combat the dissemination of SCBGAV. The miRNA-target pairs ensuring MFE exceeding the threshold standards were predicted.

As the amiRNAs have high specificity to the designed target gene, detrimental off-target effects can be minimized, permitting their silencing expression to be stably transmitted to future generations(Ossowski et al. 2008; Zhao et al. 2009). Furthermore, the small size of amiRNA permits for the insertion of multiple and distinct amiRNAs within a single gene expression cassette, which can then be transformed to develop transgenic plant resistant to multiple viruses simultaneously (Niu et al. 2006; Park et al. 2009; Schwab et al. 2010). We have designed future work to validate this promising amiRNA-based strategy can in fact be used to develop durable SCBGAV- resistance in transgenic sugarcane.

## Conclusions

The current study demonstrated a comprehensive analysis of the SCBGAV genome that is targeted by highly abundant and conserved miRNA families. This computational analysis revealed that there is high probability of targeting the SCBGAV with these conserved and abundant miRNAs against the genes coding ORF1, ORF2, and ORF3. These predicted miRNAs were identified after setting the parameters of MFE, seed pairing, target site accessibility, folding energy and pattern recognition, thus using key features of miRNA target prediction. These short-listed miRNAs are the best candidates to be utilized in sugarcane plant transformation for the development of SCBGAV-resistant sugarcane cultivars.

## Acknowledgement

This study is supported by the National Natural Science Foundation of China (No.31771865), the Sugar Crop Research System (CARS-170301) and the "Talented Young Scientist Program" (TYSP-4th Batch 2017-2018) of China Project ID (Pakistan-18-004).

## References

Adams MJ, Lefkowitz EJ, King AM, Harrach B, Harrison RL, Knowles NJ, Kropinski AM, Krupovic M, Kuhn JH, and Mushegian AR. 2016. Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2016). *Archives of virology* 161:2921-2949.

Ahmad K, Sun S-R, Chen J-L, Huang M-T, Fu H-Y, and Gao S-J. 2019. Presence of Diverse Sugarcane Bacilliform Viruses Infecting Sugarcane in China Revealed by Pairwise Sequence Comparisons and Phylogenetic Analysis. *The plant pathology journal* 35:41.

AUTREY L, BOOLELL S, JONES E, LOCIÇHART B, and NADIE A. 1992. Distribution of sugarcane Bacilliform Virus in various geographical regions. ISSCT. Proceedings XXI Congress. p 5-14.

Brodersen P, and Voinnet O. 2006. The diversity of RNA silencing pathways in plants. *TRENDS in Genetics* 22:268-280.

Dai X, Zhuang Z, and Zhao PX. 2018. psRNATarget: a plant small RNA target analysis server (2017 release). *Nucleic Acids Research* 46:W49-W54.

Doench JG, and Sharp PA. 2004. Specificity of microRNA target selection in translational repression. *Genes & development* 18:504-511.

Griffiths-Jones S, Grocock RJ, Van Dongen S, Bateman A, and Enright AJ. 2006. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Research* 34:D140-D144.

John B, Enright AJ, Aravin A, Tuschl T, Sander C, and Marks DS. 2004. Human microRNA targets. *PLoS biology* 2:e363.

Krüger J, and Rehmsmeier M. 2006. RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Research* 34:W451-W454.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, and Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome research* 19:1639-1645.

Liu S-R, Zhou J-J, Hu C-G, Wei C-L, and Zhang J-Z. 2017. MicroRNA-mediated gene silencing in plant defense and viral counter-defense. *Frontiers in Microbiology* 8:1801.

Loher P, and Rigoutsos I. 2012. Interactive exploration of RNA22 microRNA target predictions. *Bioinformatics* 28:3322-3323.

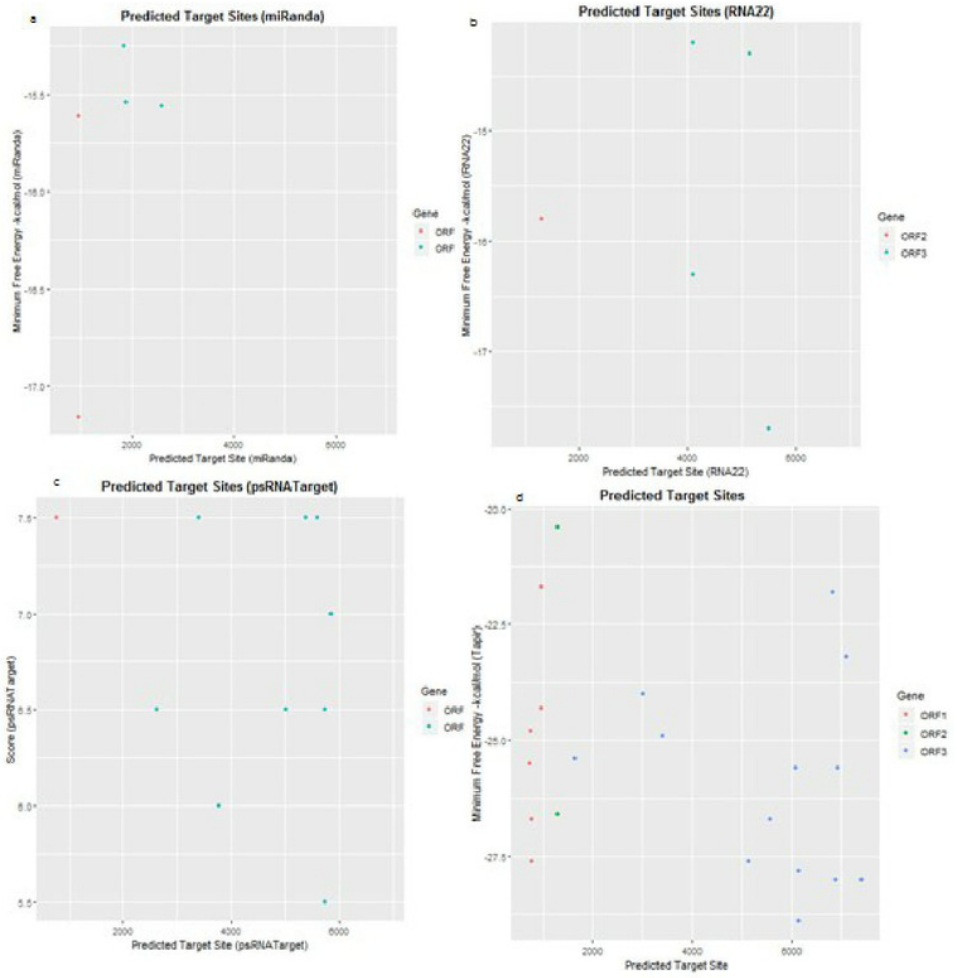
Muller E, Dupuy V, Blondin L, Bauffe F, Daugrois J-H, Nathalie L, and Iskra-Caruana M-L. 2011. High molecular variability of sugarcane bacilliform viruses in Guadeloupe implying the existence of at least three new species. *Virus research* 160:414-419.

246 Niu Q-W, Lin S-S, Reyes JL, Chen K-C, Wu H-W, Yeh S-D, and Chua N-H. 2006. Expression of
247 artificial microRNAs in transgenic Arabidopsis thaliana confers virus resistance. *Nature*
248 *biotechnology* 24:1420.
249 Ossowski S, Schwab R, and Weigel D. 2008. Gene silencing in plants using artificial microRNAs
250 and other small RNAs. *The Plant Journal* 53:674-690.
251 Park W, Zhai J, and Lee J-Y. 2009. Highly efficient gene silencing using perfect complementary
252 artificial miRNA targeting AP1 or heteromeric artificial miRNA targeting AP1 and CAL
253 genes. *Plant cell reports* 28:469-480.
254 Schwab R, Ossowski S, Warthmann N, and Weigel D. 2010. Directed gene silencing with
255 artificial microRNAs. *Plant MicroRNAs*: Springer, 71-88.
256 Sun S-R, Damaj MB, Alabi OJ, Wu X-B, Mirkov TE, Fu H-Y, Chen R-K, and Gao S-J. 2016.
257 Molecular characterization of two divergent variants of sugarcane bacilliform viruses
258 infecting sugarcane in China. *European journal of plant pathology* 145:375-384.
259 Zhao T, Wang W, Bai X, and Qi Y. 2009. Gene silencing by artificial microRNAs in
260 Chlamydomonas. *The Plant Journal* 58:157-164.
261

# Figure 1

miRNA target site locus position prediction results of SCBGAV

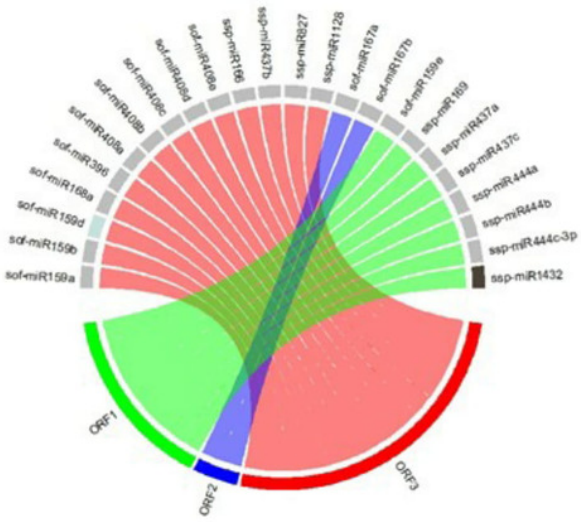
**Fig.1:** miRNA target site locus position prediction results of SCBGAV. **a)** miRanda **b)** RNA22 **c)** psRNATarget **d)** RNAhybrid.


a
Predicted Target Sites (miRanda)
Minimum Free Energy -kcal/mol (miRanda)
Predicted Target Site (miRanda)
Gene
ORF
ORF
b
Predicted Target Sites (RNA22)
Minimum Free Energy -kcal/mol (RNA22)
Predicted Target Site (RNA22)
Gene
ORF2
ORF3
c
Predicted Target Sites (psRNATarget)
Score (psRNAT arget)
Predicted Target Site (psRNATarget)
Gene
ORF
ORF
d
Predicted Target Sites
Minimum Free Energy -kcal/mol (Tapir)
Predicted Target Site
Gene
ORF1
ORF2
ORF3

# Figure 2

Circos plot showing network interaction

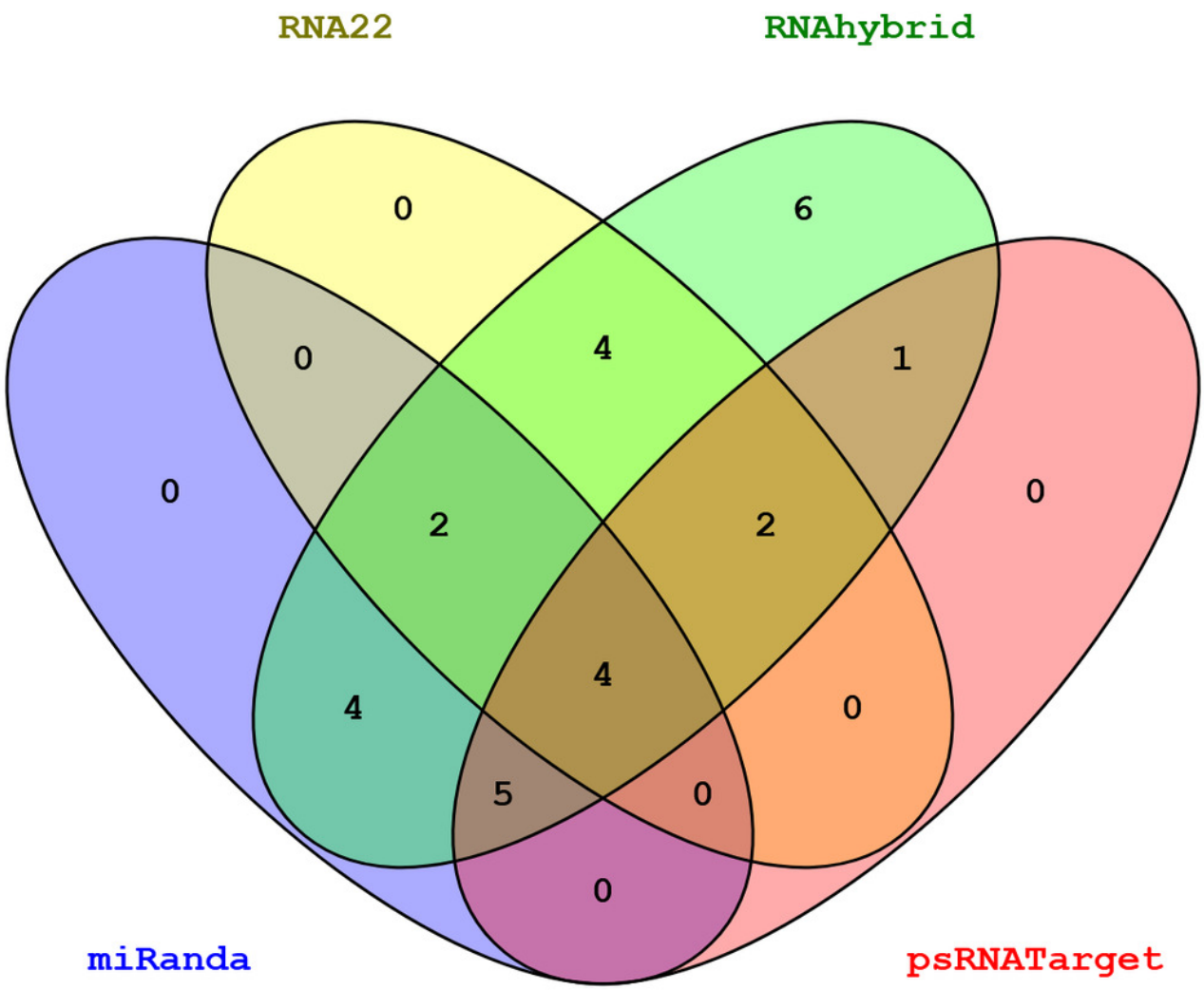
**Fig.2:** Circos plot showing network interaction between sugarcane-encoded miRNAs and their SCBGAV targets.


ssp-miR437b
ssp-miR827
ssp-miR1128
scf-miR167a
scf-miR167b
scf-miR159e
ssp-miR169
ssp-miR437a
ssp-miR437c
ssp-miR444a
ssp-miR444b
ssp-miR444c-3p
ssp-miR1432
sof-miR159a
sof-miR159b
sof-miR159d
sof-miR168a
scf-miR396i
scf-miR408a
scf-miR408b
scf-miR408c
scf-miR408d
scf-miR408e
ssp-miR166
ORF1
ORF2
ORF3

# Figure 3

Venn diagram representing common sugarcane miRNAs.

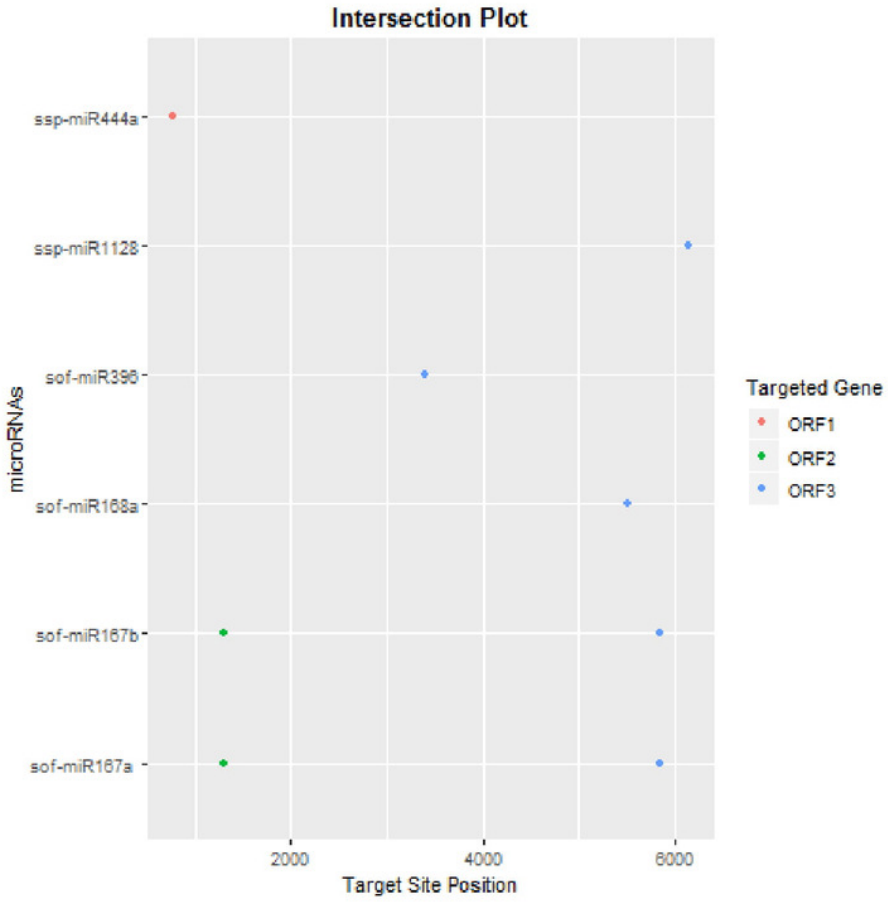
**Fig.3:** Venn diagram representing common sugarcane miRNAs predicted by all algorithms.



# Figure 4

Intersection plot showing sugarcane miRNAs

**Fig.4:** Intersection plot showing sugarcane miRNAs predicted from at least two algorithms at common position.


Intersection Plot
ssp-miR444a
ssp-miR1128
sof-miR396
sof-miR168a
sof-miR167b
sof-miR167a
microRNAs
2000
4000
6000
Target Site Position
Targeted Gene
ORF1
ORF2
ORF3

## Table 1(on next page)

miRNA target prediction algorithms

**Table1:** Characteristic features considered by miRNA target prediction algorithms under study.

1 Table.1: Characteristic features considered by miRNA target prediction algorithms under study.

| Algorithms | Main Parameters | References |
|---|---|---|
| miRanda | Seed pairing, multiple target sites and target site accessibility | John et al., 2004 |
| RNA22 | Pattern recognition, site complementarity and folding energy | Miranda et al., 2006 |
| RNAhybrid | Seed pairing, multiple target sites and target site accessibility | Rehmsmeier et al., 2004 |
| psRNATarget | Site complementarity, target site accessibility and expectation | Dai et al., 2018 |

2

## Table 2(on next page)

miRNA-target pairs were selected from miRanda analysis

**Table2:** miRNA-target pairs were selected from miRanda analysis. Locus position was selected by consensus analysis of at least two algorithms. MFE* and mode of target Inhibition** was determined by RNAhybrid.

Table.2: Selected miRNA-target pairs obtained after miRanda analysis. Locus position was selected by consensus analysis of at least two algorithms.

| **Sugarcane microRNAs** | **miRNA-target pair** | **Locus position** | **MFE* (kcal/mol)** | **Inhibition**** |
|---|---|---|---|---|
| sof-miR167(a, b) | Query: 3' gucUAGUACGACCGUCGAAGu 5'<br>Ref: 5' aaaATCAAGTT-GCAGCCTCa 3' | 5846-5865 | -26.60 | Cleavage |
| sof-mi396 | Query: 3' guCAAGUUC-UUUCGACACCUu 5'<br>Ref: 5' agGATTAGGTGATGCTGTGGAg 3' | 3394-3415 | -24.90 | Cleavage |
| ssp-miR528 | Query: 3' gaggAGACGUACGGGGAAGGu 5'<br>Ref: 5' gcgaTCCGC-CCCCCCTTCC- 3' | 7426-7444 | -28.0 | Cleavage |

MFE (Minimum Free Energy)* and mode of target Inhibition** was determined by RNAhybrid.

## Table 3(on next page)

Sugarcane miRNAs and their target positions

**Table3:** Sugarcane miRNAs and their target positions in the SCBGAV identified by algorithms. *MFE: Minimum free energy measured in /Kcal/mol where *MFE represents minimum folding energy measured in Kcal/mol.

**Table3:** Sugarcane miRNAs and their target positions in the SCBGAV identified by algorithms.

| miRNA Name | Position miRanda | Position RNA22 | Position RNAhybrid | Position psRNATarget | MFE* miRanda | MFE** RNA22 | MFE RNAhybrid | Expectation psRNATarget |
|---|---|---|---|---|---|---|---|---|
| sof-miR156 | | | 7104 | | | | -23.2 | |
| sof-miR159a | 5282 | | 1659 | 3779 | -17.18 | | -25.4 | 6 |
| sof-miR159a(1) | 6739 | | | | -18.11 | | | |
| sof-miR159b | 5282 | | 1659 | 3779 | -17.18 | | -25.4 | 6 |
| sof-miR159b(1) | 6739 | | | | -18.11 | | | |
| sof-miR159c | 6739 | | 6896 | 3779 | -16.70 | | -28 | 6 |
| sof-miR159d | 5282 | | 1659 | 3779 | -17.18 | | -25.4 | 6 |
| sof-miR159d(1) | 6739 | | | | -18.11 | | | |
| sof-miR159e | 5282 | | 733 | | -15.81 | | -25.5 | |
| sof-miR167a | 5846 | 1310 | 1304 | 5846 | -16.08 | -15.80 | -26.6 | 7 |
| sof-miR167b | 5846 | 1310 | 1304 | 5846 | -16.08 | -15.80 | -26.6 | 7 |
| sof-miR168a | 5506 | 5506 | 6084 | | -21.24 | -17.70 | -25.6 | |
| sof-miR168a(1) | 7050 | 7050 | | | -18.45 | -17.10 | | |
| sof-miR168a(2) | | 5137 | | | | -14.30 | | |
| sof-miR168b | 7050 | 7050 | 6937 | | -20.05 | -19.40 | -25.6 | |
| sof-miR168b(1) | | 5506 | | | | -13.70 | | |
| sof-miR396 | 3394 | 3394 | 3394 | 5732 | -19.99 | -17.80 | -24.9 | 5.5 |
| sof-miR396(1) | | 4115 | | | | -14.20 | | |
| sof-miR408a | | 1735 | 5136 | | | -13.70 | -27.6 | |
| sof-miR408b | | 1735 | 5136 | | | -13.70 | -27.6 | |
| sof-miR408c | | 1735 | 5136 | | | -13.70 | -27.6 | |
| sof-miR408d | | 1735 | 5136 | | | -13.70 | -27.6 | |
| sof-miR408e | | | 6152 | | | | -28.9 | |
| ssp-miR166 | 3249 | | 5572 | 981 | -21.45 | | -26.7 | 6.5 |
| ssp-miR166(1) | 5589 | | | | -17.54 | | | |
| ssp-miR166(2) | 5761 | | | | -21.74 | | | |
| ssp-miR169 | | | 752 | | | | -24.8 | |
| ssp-miR437a | | | 978 | | | | -21.7 | |
| ssp-miR437b | | | 6829 | | | | -21.8 | |
| ssp-miR437c | | | 978 | | | | -24.3 | |
| ssp-miR528 | 7426 | 1310 | 7407 | 2897 | -18.37 | -13.70 | -28 | 6.5 |
| ssp-miR528(1) | 7395 | | | | -18.05 | | | |
| ssp-miR827 | 5378 | | 3010 | | -16.40 | | -24 | |
| ssp-miR444a | | 4103 | 774 | 775 | | -16.30 | -27.6 | 6.75 |
| ssp-miR444b | | 4103 | 774 | 756 | | -16.30 | -27.6 | 6.0 |
| ssp-miR444b(1) | | | | 2615 | | | | 6.5 |
| ssp-miR444b(2) | | | | 5009 | | | | 6.5 |
| ssp-miR444b(3) | | | | 5731 | | | | 6.5 |
| ssp-miR444c-3p | | | 774 | 4940 | | | -26.7 | 6.5 |
| ssp-miR1128 | 6141 | | 6148 | | -22.96 | | -27.8 | |
| ssp-miR1432 | 6070 | | 1301 | | -15.48 | | -20.4 | |

*MFE: Minimum free energy measured in /Kcal/mol where *MFE represents minimum folding energy measured in Kcal/mol.

4