



Ab initio prediction of micro-RNA like structures in sugarcane viruses and their cellular targets

S Brindha and R Viswanathan

Plant Pathology Section, ICAR-Sugarcane Breeding Institute, Coimbatore 641007, India; rasaviswanathan@yahoo.co.in

Abstract During a viral infection in a plant, miRNAs interplay between the host and the pathogen. *Sugarcane yellow leaf virus* (SCYLV), *Sugarcane streak mosaic virus* (SCSMV), and *Sugarcane bacilliform virus* (SCBV) are serious diseases affecting sugarcane productivity worldwide. As a viable strategy to utilize resistance to the viruses in sugarcane, we used computational approaches to predict virus-encoded miRNAs in these viruses. We used an *in silico* approach to analyse SCYLV, SCSMV and SCBV virus genome sequences. Pre-miRNAs were extracted through VMir software, and screened using web tools such as Mir Para, MiPred, Mature Pred and Mature Bayes. The potential target genes were predicted using psRNATarget. From computational tools, different viral pre-miRNA hairpin sequences and 11, 8 and 13 putative mature miRNAs were predicted from SCYLV, SCSMV and SCBV genomes, respectively. These predicted miRNAs have hybridized with numerous targets. The present study is the first computational prediction of SCYLV and SCBV encoded viral miRNA and their targets in sugarcane and other monocot ESTs. This is also the first study that has identified miRNAs in P0 (SCYLV)-ORF, which is suppressor of RNA silencing. This new findings will fine-tune our strategy of using pathogen-derived resistance in sugarcane and developing virus-resistant varieties through conventional breeding.

Key words *Sugarcane yellow leaf virus*, *Sugarcane streak mosaic virus*, *Sugarcane bacilliform virus*, genomes, resistance

INTRODUCTION

In India, sugarcane productivity is significantly affected by many fungal diseases such as red rot, smut and wilt. However, recently the impact of viral diseases has been clearly visible on sugarcane yield. *Sugarcane yellow leaf virus* (SCYLV, *Polerovirus*, Luteoviridae), *Sugarcane streak mosaic virus* (SCSMV, *Poacevirus*, Potyviridae) and *Sugarcane bacilliform virus* (SCBV, *Badnavirus*, Caulimoviridae) are important viral pathogens causing yellow leaf (YL), mosaic and leaf fleck, respectively, and affect sugarcane cultivation in India (Viswanathan and Rao 2011). SCYLV infection severely affects plant growth and juice yield in susceptible varieties and is also responsible for 'varietal degeneration' in sugarcane (Viswanathan 2016; Viswanathan *et al.* 2014).

MicroRNAs (miRNAs) are a class of endogenously transcribed single-stranded non-coding small RNA species that are *21 nt in length. Recent research has revealed that the virus encoded miRNAs are the key players in modulating the antiviral host defense machinery by regulating both host cellular and their own gene expression (Song *et al.* 2011). To manage the viral disease in a long-term basis, host resistance and/or developing transgenic lines resistant to these viruses are key techniques. Here, we computationally predict the virus-specific miRNAs, which can be an effective and predictable new approach to engineering resistance to these viruses in sugarcane.

MATERIALS AND METHODS

Prediction and identification of putative pre-miRNA sequences and mature miRNAs

We used viral genome sequences of SCYLV, SCSMV and SCBV characterised in our group, including the ORFs and the 5' and 3'-UTRs of the viral RNA (GenBank accession numbers: JN925154.1, JN941985.1, and JN377535.1). The secondary structure of RNAs corresponding to each window is predicted using the MFold algorithm. We used the dedicated viral pre-miRNA prediction algorithm, VMir (Grundhoff 2006) for the initial screening with the default parameters. Further, the predicted hairpin sequences were screened using miRPara (Wu *et al.* 2011), and MiPred (Jiang *et al.* 2007) in order to discriminate real pre-miRNAs from other pseudo hairpin structures. We used MaturePred (Xuan *et al.* 2011) and MatureBayes (Gkirtzou *et al.* 2010) web tool to extract mature miRNA: miRNA* duplexes from pre-miRNA hairpins.



Prediction of secondary structure and potential targets

To predict the secondary structure of pre-miRNA, we used the Zuker RNA folding algorithm software MFOLD 3.1 (Zuker 2003) to generate RNA secondary structures and calculate the minimum free energy. To predict the potential target genes by SCYLV, SCSMV and SCBV-encoded miRNAs, we used the psRNATarget (<http://plantgrn.noble.org/psRNATarget/>) against the pre-loaded transcript/genomic library of datasets present in the library of psRNATarget server.

RESULTS AND DISCUSSION

Using an *ab initio* approach, we identified 11 SCYLV, 8 SCSMV, 13 SCBV putative mature miRNAs (Figs 1-3). These miRNAs were dispersed in both UTR, ORFs and intergenic regions of the viral genomes. The length of the pre-miRNAs falls with an average length of nucleotides, which is in the same range as the pre-miRNA experimentally identified. The diverse lengths of the identified miRNAs suggest unique roles for modulation of miRNA biogenesis or gene expression (Pani *et al.* 2011). Remarkably, some of the mature miRNAs start with a 5'-terminal uridine residue, a typical characteristic of miRNAs recognized by the AGO1 protein. These findings indicate that the predicted miRNAs may be noncanonical (Wan *et al.* 2012). Among the 11 SCYLV miRNA, four of the miRNA are located in the ORFs P0, which is suppressor of RNA silencing, and pathogenesis-related proteins targeting pathways of functionally important genes for plant defense.

Prediction of the potential targets for putative miRNAs in sugarcane viruses

It is necessary to identify the targets of these viral encoded miRNAs to understand the biologic functions of these miRNAs in pathogen infection. The predicted viral miR/miRs* showed complementarity to several regions in the *Arabidopsis*, *Oryza*, *Zea*, *Sorghum* and *Saccharum* genomes, including proteins such as leaf senescence protein, F-box protein, regulator of chromosome condensation (RCC1) family protein, zinc finger protein (PRAF1), jasmonate inducible protein, chloroplast chlorophyll a/b binding protein, cold-regulated protein cor15b precursor, heat shock cognate protein 70, serine/threonine kinases, Myb, zinc finger proteins and MADS-box transcription factor. All these proteins have their functions related to morphogenesis, disease resistance, migration, biotic response, metabolic pathways, transcription factors and growth pathways.

CONCLUSIONS

Computational approaches allowed us for the first time to show that SCYLV and SCBV encode possible pre-miRNA hairpins, and the expression of these predicted miRNAs could be detected using stem loop-RT PCR. The method led to the identification of virus miRNAs encoded in ORFs P0 of SCYLV, which are suppressors of RNA silencing. These findings can be further used to engineer virus resistance via RNA-based strategies, which could offer a long-term disease solution by providing resistance to the important sugarcane viruses.

ACKNOWLEDGEMENTS

This study was supported by the Department of Biotechnology, New Delhi under RNAi-BT/PR4978/AGR/36/712/2012.

REFERENCES

- Gkirtzou K, Tsamardinos I, Tsakalides P, Poirazi P. 2010. MatureBayes: a probabilistic algorithm for identifying the mature miRNA within novel precursors. *PLoS One* 5: e11843.
- Grundhoff A, Sullivan CS, Ganem D. 2006. A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. *RNA* 12: 733-750.
- Jiang P, Wu H, Wang W, *et al.* 2007. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic Acid Research* 35: 339-344.
- Pani A, Mahapatra RK, Behera N, Naik PK. 2011. Computational identification of sweet wormwood (*Artemisia annua*) microRNA and their mRNA targets. *Genomics Proteomics Bioinformatics* 9: 200-210.
- Song L, Gao S, Jiang W, *et al.* 2011. Silencing suppressors: viral weapons for countering host cell defenses. *Protein Cell* 2: 273-281.
- Viswanathan R. 2016. Varietal degeneration in sugarcane and its management in India. *Sugar Tech* 18:1-7.



- Viswanathan R, Chinnaraja C, Malathi P, *et al.* 2014. Impact of *Sugarcane yellow leaf virus* (SCYLV) infection on physiological efficiency and growth parameters of sugarcane under tropical climatic conditions in India. *Acta Physiologiae Plantarum* 36: 1805-1822.
- Viswanathan R, Rao GP. 2011. Disease scenario and management of major sugarcane diseases in India. *Sugar Tech* 13: 336-353.
- Wan LC, Zhang H, Lu S, *et al.* 2012 Transcriptome-wide identification and characterization of miRNAs from *Pinus densata*. *BMC Genomics* 13: 132.
- Wu Y, Wei B, Liu H, *et al.* 2011. MiRPara: a SVM-based software tool for prediction of most probable microRNA coding regions in genome scale sequences. *BMC Bioinformatics* 12: 107.
- Xuan P, Guo M, Huang Y, Li W, Huang Y. 2011. MaturePred: efficient identification of microRNAs within novel plant pre-miRNAs. *PLoS One* 6: e27422.
- Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* 31: 3406-3434.

Prédiction *ab initio* de structures ressemblant à des micro-ARN dans des virus de la canne à sucre et leurs cibles cellulaires

Résumé. Au cours de l'infection virale d'une plante, des miARN interagissent avec l'hôte et l'agent pathogène. Le *Sugarcane yellow leaf virus* (SCYLV), le *Sugarcane streak mosaic virus* (SCSMV) et le *Sugarcane bacilliform virus* (SCBV) provoquent des maladies graves affectant la productivité de la canne à sucre dans le monde. Nous avons utilisé des approches informatiques afin de prédire les miARN potentiellement codés par ces virus en tant que stratégie viable permettant d'utiliser la résistance aux virus chez la canne à sucre. Nous avons utilisé une approche *in silico* pour analyser les séquences de génome viral du SCYLV, du SCSMV et du SCBV. Des pré-miARN ont été extraits en utilisant le logiciel Vmir, puis criblés en utilisant des outils d'internet tels que Mir Para, MiPred, Mature Pred and Mature Bayes. Les gènes cibles potentiels ont été prédits grâce à psRNATarget. Les outils informatiques ont permis de prédire différentes séquences virales en épingle à cheveux de pré-miARN ainsi que 11, 8 et 13 mi-ARN putatifs matures des génomes de respectivement SCYLV, SCSMV et SCBV. Ces miARN prédits se sont hybridés avec de nombreuses cibles. La présente étude est la première à avoir permis la prédiction informatique de miARN viraux codés par les virus SCYLV et SCBV ainsi que leurs cibles dans la canne à sucre et d'autres EST de monocotylédones. C'est aussi la première étude à avoir identifié des miARN dans l'ORF PO de SCYLV qui est suppresseur d'extinction d'ARN. Ces nouvelles découvertes permettront de perfectionner notre stratégie d'utilisation de la résistance dérivée du pathogène chez la canne à sucre et de développement de variétés résistantes aux virus par création variétale conventionnelle.

Mots-clés: *Sugarcane yellow leaf virus*, *Sugarcane streak mosaic virus*, *Sugarcane bacilliform virus*, genomes, resistance

Modelado *ab initio* de estructuras similares a micro-ARN en virus de la caña de azúcar y sus objetivos celulares

Resumen. Durante la infección viral de la planta, los miRNAs interactúan entre el hospedero y el patógeno. Los virus *Sugarcane yellow leaf virus* (SCYLV), *Sugarcane streak mosaic virus* (SCSMV), y *Sugarcane bacilliform virus* (SCBV) son patógenos graves que afectan la productividad de la caña de azúcar en todo el mundo. Como una estrategia viable para utilizar la resistencia a virus en caña de azúcar, se utilizó métodos computacionales para predecir miRNAs codificadas por estos virus. Se utilizó un enfoque *in silico* para analizar las secuencias del genoma del virus SCYLV, SCSMV y ScBV. Pre-miRNAs se obtuvieron a través del software Vmir, y se evaluaron usando herramientas web como Mir Para, MiPred, Mature Pred y Mature Bayes. Los genes objeto de estudio y potenciales se predijeron usando psRNATarget. Con herramientas computacionales se predijeron diferentes secuencias virales pre-miARN tipo horquilla, y 11, 8 y 13 posibles miARNs maduros, a partir de los genomas del SCYLV, SCSMV y ScBV, respectivamente. Estos miARNs identificados hibridaron con numerosas secuencias objetivo. Este es el primer estudio que predice computacionalmente miARN virales codificados en SCYLV y SCBV, y sus secuencias objetivo en caña de azúcar y otros EST de monocotiledóneas. Este también es el primer estudio que ha identificado miARNs en PO(SCYLV)-ORF, que es supresor del silenciamiento de ARN. Estos nuevos hallazgos afinarán nuestra estrategia para usar la resistencia derivada del patógeno en caña de azúcar, y el desarrollo de variedades resistentes a virus a través del mejoramiento convencional.

Palabras clave: *Sugarcane yellow leaf virus*, *Sugarcane streak mosaic virus*, *Sugarcane bacilliform virus*, genomas, resistencia

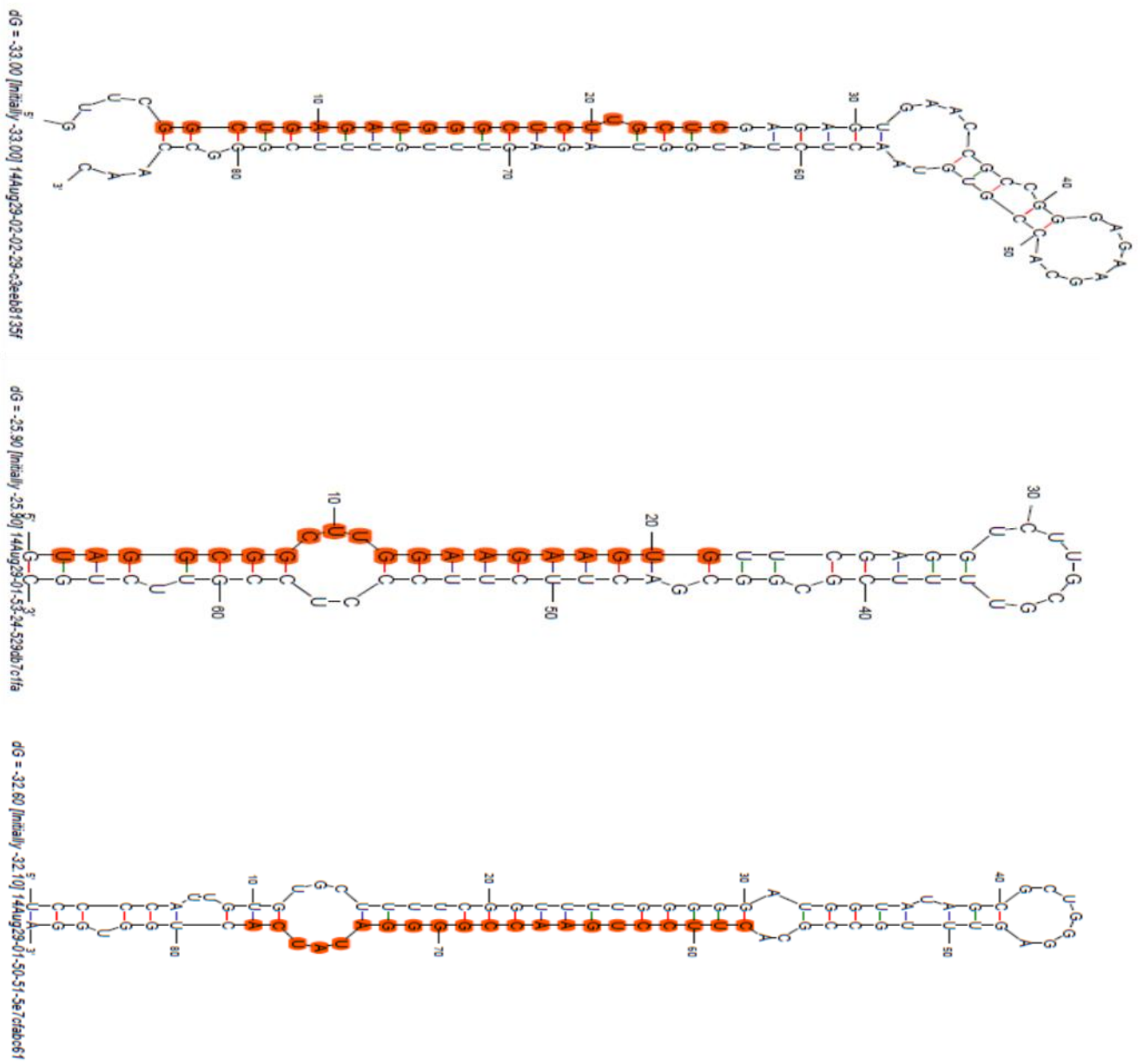


Fig. 1. Predicted pre-miRNA secondary structures for *Sugarcane yellow leaf virus* (SCYLV). Mature miRNA sequences are given in red.



Fig. 2. Predicted pre-miRNA secondary structures for *Sugarcane streak mosaic virus* (SCSMV). Mature miRNA sequences are given in red.

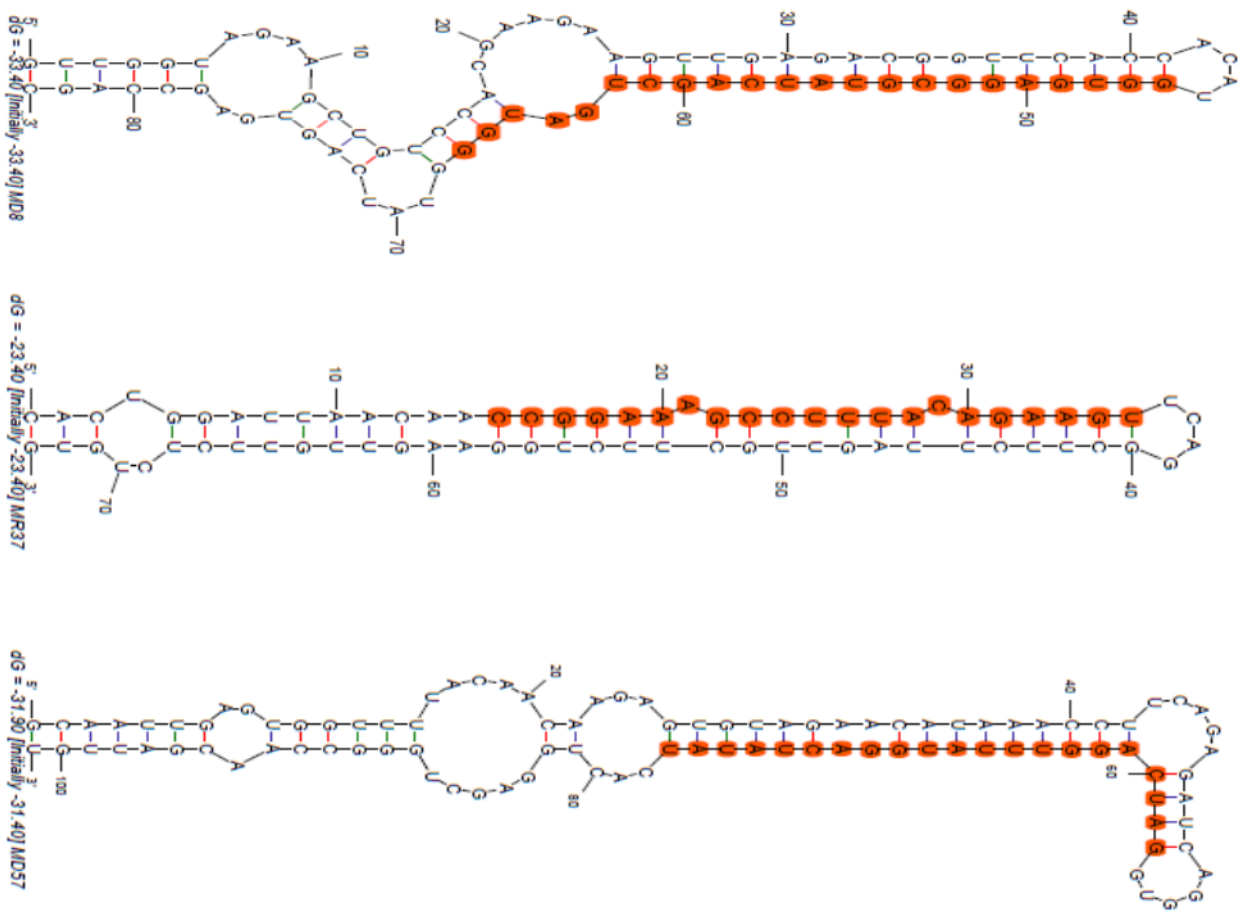


Fig. 3. Predicted pre-miRNA secondary structures for *Sugarcane bacilliform virus* (SCBV). Mature miRNA sequences are given in red.