

Prediction and Identification of MicroRNA from Banana Infected with Banana Streak Mysore Virus (BSMYV)

M. Mary Sheeba, R. Selvarajan* and M.M. Mustaffa

Molecular Virology Lab National Research Centre for Banana Tiruchirapalli, Tamil Nadu, India

In this study, computational prediction of miRNAs from banana EST and BAC sequences was performed, which resulted in the finding of 18 banana miRNAs belonging to 16 miRNA families. For the 18 miRNAs, 25 banana ESTs were found as their potential targets. Most of these targets encode transcriptional factors which play a role in banana development. Additionally, real time PCR assays were performed to profile the expression levels of three miRNAs (miR156, 159 and 166) after the infection of *Banana streak Mysore virus* (BSMYV). The results showed that symptom severity is correlated to the expression level of miRNAs, and a clear increased expression of miR166 during viral infection. This is the first report to predict and identify miRNAs in banana.

Key words: Banana; EST; MicroRNA; RT-PCR; banana streak mysore virus (BSMYV).

*Corresponding author email: selvarajanr@gmail.com

MicroRNAs (miRNAs) are a class of highly evolutionary conserved endogenous non- coding small RNAs with about ~21 nucleotides in length (Ambros, 2001; Ambros *et al.*, 2003). It is a highly specific regulator of gene expression and derives from larger precursors that are transcribed from non-coding RNAs (Bartel, 2004). Plant miRNAs generally interact with their targets through perfect or near-perfect complementarity and cause transcriptional repression or target mRNA degradation (Llave *et al.*, 2002; Rhoades *et al.*, 2002; Bartel, 2004).

Computational prediction of miRNAs is an effective way to identify a large number of miRNAs from different plants and animals. Among different computational strategy, the homology search-based approach is an efficient strategy for identifying plant miRNAs using Expressed Sequenced Tag (EST) analysis (Zhang et al., 2005). Using computational strategies, a total of 16772 hairpin precursor miRNAs and 19724 mature miRNAs were identified from 153 species, including animals, plants and viruses (miRBase Release 17, 2011, http:// microrna.sanger.ac.uk/sequences/) (Kozomara and Griffiths- Jones, 2011). In plants, a total of 8433 miRNAs has been identified from Arabidopsis thaliana, Brassica napus, Glycine max, Medicago truncatula, Oryza sativa, Physcomitrella patens, Populus trichocarpa, Saccharum officinarum, Sorghum bicolor and Zea mays, distributing in 121 miRNA families (Zhang et al,. 2010).

In the course of co-evolution with host plants, viruses have also developed counter defense strategy to escape the defensive systems of host

plants (Simo'n-Mateo and Garcý', 2006). Sequence specific RNA degradation pathway directed by small interfering RNAs (siRNAs) restrict the accumulation and spread of exogenous virus invaders (Mlotshwa *et al.*, 2008). To overcome this strategy, most plant viruses have evolved suppressor proteins to counteract host RNA silencing (Chapman *et al.*, 2004). Similar to siRNAdirected RNA degradation, miRNA metabolism can also be altered by the activity of viral silencing suppressors through attacking common elements of the two pathways (Bazzini *et al.*, 2007). Also, plant can generate miRNAs during viral infection, which are involved in the regulation of the virus defense process in plants or targeting some key genes of virus development to suppress its increase (Lu *et al.*, 2008).

Banana (Musa spp.) is one of the most importantly economical crops in India. Prediction of miRNAs is poorly understood in spite of its large genome (600Mb). Hence this study was initiated to efficaciously search the conserved banana miRNA homologues from the publicly available Genomic Survey Sequences (GSS), EST, Core Nucleotide databases and Bacterial Artificial Chromosomes (BAC) sequences. First, we identified 18 conserved miRNAs and detected 25 potential targeted genes. Additionally, real-time PCR assays were performed to profile the expression levels of 3 miRNAs (miR156, 159 and 166) after the infection of Banana streak Mysore virus. The results showed that symptom severity is correlated to the miRNA accumulation, and increased expression of all three miRNAs during virus infection. Furthermore, miR166 expression level in banana leaves was also surveyed after the Banana Streak Mysore virus

(BSMYV) infection, which is expected to yield the greater understanding of the involvement of miRNA in viral infection and leaf patterning. This paves a way to understand the plant-pathogen interactions and host defense signaling pathways.

Materials and Methods

Prediction of potential miRNAs

To search miRNA precursors, a total of 44114 banana nucleotide sequences was downloaded from NCBI GenBank nucleotide databases, including 5179 in Core Nucleotide, 31745 in EST and 7190 in GSS (2009). Additionally, 49 BACs (finished BACs 2009) sequences were obtained from GNPAnnot Portal sequencing data (2009; http://www.gnpannot. org/fr/content/musaceae - stastistics). The method used in the study was described by Zhang and coworkers (Zhang et al., 2006). To obtain the hits for sequences with the perfect match against known miRNAs, previously reported banana sequences (Core Nucleotide, GSS, EST and BACs sequences) were searched with all known non-redundant mature miRNA sequences identified or predicted from other plants. The repeat sequences and those coding proteins were then removed from the hit sequence. BLAST 2.2.4 (ftp://ftp.ncbi.nih.gov/blast/) and BLASTX (http://www.ncbi.nlm.nih.gov/BLAST/) (Altschul et al., 1997) were used for the searching sequences with the perfect match and non-coding proteins. The secondary structures of candidate sequences were predicted using the web-based software Mfold 3.2 (http://frontend.bioinfo.rpi.edu/applications/mfold/ cgibin/rna-form1.cgi) (Zuker, 2003). The potential premiRNAs should meet all the following criteria:

a) the secondary structure of the candidate sequences should have the stem-loop structure and no loop or break in the mature miRNA sequences; b) miRNA should have less than six mismatches with the antisense miRNA (miRNA*) sequences; c) the mature miRNA and miRNA* site must be in one arm of the stem-loop structure; d) the candidate pre-miRNA sequence should have 30–75% A+U contents; e) potential precursors have a higher minimal folding free energy (MFE) index (MFEI) than other types of RNAs. The MFEI was calculated using the following equation: MFEI = [(MFE/ length of the RNA sequence) × 100] / (G +C) %

Where MFE denotes the negative folding free energies (ÄG); Predicted miRNAs and their related information were recorded. Closely related EST and GSS sequences were blasted against each other and analyzed.

Prediction of potential target genes

A computational approach was employed based on the previous studies that all miRNAs regulate gene expression by binding to targeted mRNA sequences in a perfect or near perfect complementary site (Bartel, 2004). The number of allowed mismatches at complementary sites between miRNA sequences and potential mRNA targets was four or fewer and no gaps were allowed at the complementary sites. We utilized miRU software to predict the complementary miRNA and the target sites with default parameters (Zhang, 2005). We also used the identified banana miRNAs to do a BLASTn search against the EST databases with an E-value cutoff of 1.0 for a similarity search.

Virus inoculation

Tissue culture banana plants (*Musa balbisiana* Subgroup Mysore Syn: Poovan) were grown under glasshouse conditions with a 16h photoperiod at 23-28_°C. BSMYV was maintained on banana plant, and inoculated to the tissue culture plants by mealy bugs. At 21 days post-inoculation (dpi), the upper systemically leaves were harvested.

Quantification of miRNAs by qRT-PCR

Total RNA was isolated from virus inoculated and healthy leaves using LiCl method (Chang et al., 1993). One micro gram of total RNA was mixed with 1 µl of 10X reaction buffer with MgCl₂ and 1.0U of DNase. The mixture was incubated @ 37°C for 30 min. finally 1.0 of 25mM EDTA was added and incubated @65°C for 10 min to inactivate DNase enzyme. Fifty nanograms of total RNA was mixed with 0.5µg of oligo dT with adaptor (CGAACATGTACAGTCCATGGATAG d(T)30). The mixture was incubated at 65°C for 5min, chilled on ice for 2 min and 7µl of RT mixture containing 50mM Tris-HCl, pH 7.3, 75mM KCl, 10mM DTT, 2.5mM MgCl₂, 0.5-2mM dNTP, 20U Ribolock RNase inhibitor and 200U of MMuLV-RT enzyme (MBI Fermentas, UK) was added and the final volume was made up to 20 µl by adding sterile nuclease -free water. Samples were incubated at 42°C for 60 min and subsequently incubated at 70°C for 5 min. For miRNAs quantification, their mature sequences were downloaded from miRNA registry database, and forward and reverse primers were designed. The miRNA specific forward primers used in this study are as follows: miR-156 TGACAGAAGAGAGTGAGCACA; miR-159 TTTGGATTGAAGGGAGCTCTA; miR-166 CGGACCAGGCTTCATTCCCC; miR-169 TAGCCAAGGATGACTTGCCTG and the common primer reverse adaptor was CGAACATGTACAGTCCATGGATAG. The banana beta-actin primers (forward 5' -GGTAT GGTGT TGGAT TCTGG A-3'; reverse 5 '-GTAGT CTCAT GGATA CCTGC-3') was chosen as the reference endogenous gene. All the primer pairs were optimized and validated. PCR volumes were set up to 20 µl that contained 10 µl 2X Power SYBR Green PCR master mix (Applied Biosystems), 5 µl of a 1:1 dilution of the cDNA template, and 200nM each of the corresponding forward and reverse primer. The cycling profile was 95°C for 10s, followed by 40 cycles of 10s at 95°C and 1min at 60°C. all the

reactions were performed in triplicate, and the control with no template was included for each gene. Immediately after the final PCR cycle, a melting cycle analysis was done to determine the specificity of each reaction. The threshold cycle (C_T) values were determined automatically by Applied Biosystem's Stepone Plus system, and the fold changes of each gene were calculated as relative quantity (RQ) values using the comparative Ct (2—ct) method. The resulting PCR products were cloned into pGEM-T easy vector (Promega) and sequenced. The sequences were subsequently processed and

vector sequences were removed using the VecScreen program. The trimmed sequences were blasted against known miRNAs in the miRBase database. Candidates with perfect matches were used to fold-back secondary structure prediction with the Mfold program. The target prediction was performed as per the previous section.

Results and Discussion

Identification of Potential banana miRNAs

The first step was to search for potential hairpin structures in the banana ESTs sequences which

Table 1. List of miRNAs identified from ESTs and BAC clone	es of banana
--	--------------

miRNA family	miRNA sequences	Gene ID	Gene Source	Location	NM(Nt)	LM(Nt)	LP(Nt)	GC%	MFEkcal/mole	MFEI
156c	tgacagaagagagtgagcaca	146226345	EST	5'	0	21	98	56.1	-58.3	1.06
156h	tgacagaagagagagagcat	146223567	EST	3'	0	20	125	54.4	-34.50	0.51
159	tttggattgaagggagctcta	197650521	EST	5'	0	21	218	48.2	-93.5	0.89
160	tgcctggctccctgcatgcca	197653670	EST	5'	0	21	125	56	-50.4	0.72
164	tggagaagcaggtcacgtgca	189011744	BAC	3'	1	21	91	58.2	-30.0	0.56
166	tcggaccaggcttcattcccc	182664514	EST	5'	0	21	88	56.8	-43.5	0.87
169	tagccaaggatgacttgcctg	197637388	EST	5'	0	21	129	54.3	-62.5	0.89
396c	ttccacagctttcttgaactt	189011743	Nr	5'	0	21	101	42.6	-22.6	0.53
396f	ctccacaggctttcttgaactg	189011742	BAC	5'	0	22	111	43.2	-44.50	0.92
397	tcattgagtgcagcgttgatg	197651606	EST	5'	0	21	131	48.9	-58.0	0.91
399	tgccaaaggagaattgccctg	197654832	EST	5'	0	21	92	54.3	-48.6	0.97
444	ttgctgcctcaagcttgctgc	197655404	EST	3'	0	21	91	47.3	-21.6	0.50
845	cggctctgataccaattgatg	102139931	BAC	3'	0	21	151	28.5	-32.8	0.76
1310	ggcatcggggggcgcaacgccc	146222782	EST	5'	3	23	79	59.5	-30.9	0.65
2083	ttcttgcactcctccatctct	197637552	EST	3'	1	21	106	59.4	-35.5	0.56
2118	ttgccgattcctcccatcccta	197654740	EST	5'	2	22	137	59.9	-81.2	0.98
2914	catggtggtgacgggtgacggag	197638731	EST	5'	0	23	63	55.6	-21.8	0.62
2916	tggggactcgaagacgatcatat	146222830	EST	3'	4	23	91	57.1	-25.9	0.50

NM: number of mismatch; LM: length of mature miRNAs; LP: length of precursor; MFEs: minimal folding free energies; MFEIs: minimal folding free energy index; Nt: Nucleotide; Nr:

yielded 63 miRNA homologs. Within the 63 miRNAs predicted, 57 miRNAs were identified in 31745 EST search and 9 miRNAs in BAC search. The second step was to search for pre-miRNA based on their



nucleotide composition and free energy of the secondary structure which yielded 18 pre-miRNAs. Out of 18 miRNAs, 14 miRNAs were identified in the EST database and 4 miRNAs in the BAC database (Fig.1 a-d; Table 1).





Fig. 1a, b, c, and d Predicted precursor miRNA structures of banana miRNAs identified in this study.Mature miRNA sequences are underlined.Expression analysis of banana miRNAs during
viral infectionaxillary meristem differentiation, and leaf morphology.
Therefore, the expression alterations of these miRNAs

MiR156 and miR159 are reported to be implicated in promoting floral transitions and regulate flowering time. MiR166 is essential regulators of meristem initiation and maintainence, axillary meristem differentiation, and leaf morphology. Therefore, the expression alterations of these miRNAs and target mRNAs upon virus infections were determined by quantitative real time-PCR in this study. After BSMYV inoculation, the disease symptoms were monitored between 7 and





35dpi. BSMYV infection severely altered leaf morphogenesis, inducing the chlorotic streak and necrosis in contrast to healthy controls. To investigate the interference of viral infections with miRNA pathways in banana, the expression patterns of three miRNAs were quantified at 21dpi. The amount of miRNA species in healthy mock plants was arbitrarily set at 1.0, and other data were

Table 2. List of the potential targets of conserved banana miRNAs

miRNA Family	Target gene	Conserved for putative target gene	Conserved for Oryza sativa target gene
156	1. Squamosa promoter binding protein (SBP) DNA binding/ Transcription factor	ES434709	Os06g49010
	2. Phospholipase A1	-	Os09g33820
	3. Teosinte glume architecture	FF558536	Os11g30370
450		F I 000004	0.04.50000
159	1. Myb family transcription factor (MYB120), GAMYB	FL666961	0501959660
	2. Ubiquitin-protein ligase E3C	FL661284	Os03g47949
	3. rac GTPase activating protein 1	FL662243	Os12g05900
		FL660736	
160	1. Auxin-responsive factor (ARF16)	FL666054	Os06g47150, Os04g43910,
		DN238918	Os02g41800
	2. Sugar transport protein 5		Os04g37990
164	1. NAC1/NAM protein (Cup-shaped cotyledon 1 protein)	-	Os08g10080, Os02g36880
166	1. Homeobox-leucine zipper transcription factor	FF558732	Os12g41860, Os03g43930
	2. Retrotransposon protein		Os07g08900, Os11g45899
	3. Rolled leaf1, putative		Os03g01890
169	1. CCAAT-binding transcription factor (CBF-B/NF-YA)	FL648124 FL646383 FL646729 FL648568	Os12g42400, Os03g07880 Os03g29760, Os03g48970
396	1. Ubiquitin-protein ligase COP1		Os01g52640
	2. Jasmonate O-methyltransferase	-	Os04g57050, Os04g57070
	3. Heat shock 70 kDa protein 4		Os01g08560
397	 L-ascorbate oxidase precursor monocopper oxidase-like protein SKS1 precursor 	FL666615 FL663510 FL661805	Os01g62490, Os01g63200 Os11g48060
399	1. Phosphate transporter	FL666459	Os08g45000
	2. Ubiquitin conjugating enzyme family protein		Os05g48390
444	1. MADS box protein 2	FL668327	Os02g49840, Os08g33488
	2. Chitin-inducible gibberellin-responsive protein 2		Os07g39470,
	3. Retrotransposon protein, putative, Ty3-gypsy subclass		Os05g07580, Os01g64070
845	1. Retrotransposon protein	FF562629	Os01g65760, Os09g37850
			Os05g22390, Os01g03230
1310	1. Cyanogenic beta-glucosidase precursor	-	Os11g08120
	2. Protein binding protein		Os03g55950, ES432023
2083	1. OsWAK33 - OsWAK receptor-like protein		Os04g21820, Os04g39570
	2. OsIAA8 - Auxin-responsive Aux/IAA gene family member	FL648759	Os02g49160, Os01g18360
	3. Ribulose bisphosphate carboxylase	FL648906	
		FL648903	
2118	1. NB-ARC domain containing protein	FL668165	Os08g24380
	2. Disease resistance protein(CC-NBS-LRR)	FL664303	Os06g30430,Os08g43010
2914	1. F-box domain containing protein	-	Os11g37060, Os08g16710
2916	1. F-box domain containing protein	-	Os05g25580

computed relative to these plants as shown in Fig.2. Compared with mock inoculated plants, BSMYV induced a clear increase in miR166 levels (RQ=31.3). MiR156 and miR159 were also increased in BSMYV infected plants to levels of RQ>5. Comparing with the virus symptoms on infected plants, we found that miRNA accumulated to a greater abundance in the virus that produced severe symptom. Comparing with the virus symptoms on infected plants, we found that miRNA accumulated to a greater abundance in the virus that produced severe symptom. Comparing with the virus symptoms on infected plants, we found that miRNA accumulated to a greater abundance in the virus that produced severe symptom (BSMYV). It has previously been shown that PVX causes very mild symptoms in *N. benthamiana* and also causes only moderate changes in the accumulation of miR156,

miR159, miR160, miR164, miR165, miR166, miR167 and miR169 (Bazzini *et al.*, 2007), similarly the accumulation of miR159/319 and miR172 was observed to increase with the days post inoculation of *Tomato leaf curl New Delhi virus* (ToLCNDV) in tomato cv Pusa Ruby (Naqvi *et al.*, 2010) and the same trend of increase in miRNA levels has been reported with the begomovirus infection (Amin *et al.*,2011)

Similarly, in CMV-Fny and CMV-Fny-satT1 infected plants, miR165/166 was increased to level of RQ> 7.0 and miR159 was also increased to level of approximately RQ=5.0 at 21dpi (Feng *et al.*, 2012). The PCR amplified fragments of size ~110bp were cloned and sequenced. The sequences were processed and trimmed sequences were folded to

form secondary structure using Mfold program. These amplified miRNAs were obtained as that of predicted sequences.

Prediction of miRNAs targets in Banana

To identify potential regulatory targets, we first searched for banana mRNAs in database that were complementary, with less than four mismatches, to the 18 miRNAs. Gaps, G: U and other noncanonical pairs were not allowed and considered as mismatches according to the screening criteria. Using the 18 newly identified miRNAs against mRNA/EST sequences of banana, we found 25 target genes (Table 2). The different targets can be separated into several groups. The first group was predicted to encode transcription factors. The next group contains miRNA targets coding a range of different proteins which may play important roles in the aspects of metabolisms, development, signal transduction, and stress response. However, the functions of some targets were largely unknown. In our study, it appears that our predicted targets play roles not only in development, but also in diverse physiological processes.

In summary, in the present study, we have predicted conserved banana miRNAs and their targets. A total of 18 potential miRNAs, belonging to 16 miRNA families, and 25 of their potential target genes were identified. Most of the pooled targets were predicted with functions in a variety of biological processes, including growth and developmental patterning, metabolic processes, hormone responses, biotic stress defense and signaling. Among biotic stress defense, miRNAs has long been employed as a strategy to resist viruses. This study is a foremost step to characterize the miRNAs which are probably involved in the plant viral defense mechanisms. Further mechanisms of miRNA biogenesis and evolution vis-à-vis virus infection in plants are yet to be understood to elucidate the complicated roles in viral infection networks. Knowing the function of miRNA in plant-virus interaction will pave the way for miRNA mediated anti-virus defense.

Acknowledgements

We thank the Director, NRCB, Tiruchirapalli for providing facilities.

References

- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.*, 25: 3389–3402.
- Ambros, V. 2001. MicroRNAs: tiny regulators with great potential. *Cell*, **107**: 823–826.
- Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, D., Eddy, S.R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G. and

Tuschl, T. 2003. A uniform system for microRNA annotation. *RNA*, **9**: 277–279.

- Amin, I., Patil, B.L., Briddon, R.W. Mansoor S. and Fauquet, C.M. 2011. A common set of developmental miRNAs are upregulated in *Nicotiana benthamiana* by diverse begomoviruses. *Virol. J.*, 8:143.
- Bartel, D. 2004. MicroRNAs: genomics, biogenesis, mechanism and function. Cell, 116: 281–297.
- Bazzini, A.A., Hopp, H.E., Beachy, R.N. and Asurmendi, S. 2007. Infection and co- accumulation of tobacco mosaic virus proteins alter microRNA levels, correlating with symptom and plant development. *Pro. Natl. Acad. Sci. USA*, **104**: 12157-12162.
- Chang, S., Puryear, J. and Cairney, J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.*, **11**: 113-116.
- Chapman, E.J., Prokhnevsky, A.I., Gopinath, K., Dolja, V.V. and Carrington, J.C. 2004. Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. *Genes Development*, **18**: 1179– 1186
- Feng, J., Lai, L., Lin, R., Jin, C., and Chen, J. 2012. Differential effects of cucumber mosaic virus satellite RNAs in the perturbation of microRNA-regulated gene expression in tomato. *Mol. Biol. Rep.*, **39**: 775–784.
- Kozomara, A.and Griffiths-Jones, S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.*, **39**: D152-D157.
- Llave, C., Kasschau, K.D., Rector, M.A., and Carrington, J.C. 2002. Endogenous and silencing-associated small RNAs in plants. *Plant Cell*, **14**: 1605–1619.
- Lu, Y.D., Gan, Q.H., Chi, X.Y., and Qin, S. 2008. Roles of miRNA in plant defense and virus offense interaction. *Plant Cell Rep.*, 27: 1571-1579.
- Mlotshwa, S., Pruss, G.J. and Vance, V. 2008. Small RNAs in viral infection and host defense. *Trends Plant Sci.*, **13**: 375–382.
- Naqvi A.R., Haq, Q.M.R. and Mukherjee, S.K. 2010. MicroRNA profiling of tomato leaf curl new Delhi virus (tolcndv) infected tomato leaves indicates that deregulation of mir159/319 and mir172 might be linked with leaf curl disease. *Virol J.*, **7:**281
- Rhoades, M.W., Reinhart, B.J., Lim, L.P., Burge, C.B., Bartel, B. and Bartel, D.P. 2002. Prediction of Plant MicroRNA Targets. *Cell*, **110**: 513–520.
- Simo'n-Mateo, C. and Garcya, J.A. 2006. MicroRNA guided processing impairs Plum pox virus replication, but the virus readily evolves to escape this silencing mechanism. *J. Virol.*, **80**: 2429–2436.
- Zhang, B.H., Pan, X.P., Wang, Q.L., Cobb, G.P. and Anderson, T.A. 2005. Identification and characterization of new plant microRNAs using EST analysis. *Cell Res.*, **15**: 336–360.
- Zhang, B.H., Pan, X.P., Wang, Q.L., Cobb, G.P. and Anderson, T.A. 2006. Computational identification of microRNAs and their targets. *Comp. Biol. Chem.*, **30**: 395–407.
- Zhang, Z., Yu, J., Li, D., Zhang, Z., Liu, F., Zhou, X., Wang, T., Ling, Y. and Su, Z. 2010. PMRD: Plant microRNA database, *Nucleic Acids Res.*, **38**: D806-D813.
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**: 3406–3415.

Received: September 14, 2012; Accepted: March 1, 2013