Journal of Root Crops, 2020, Vol. 46 No. 1, pp. 41-45
Indian Society for Root Crops
ISSN 0378-2409, ISSN 2454-9053 (online)

Identification of Cassava (*Manihot esculenta* Crantz.) miRNAs Targeting the Genome of Cassava Mosaic Virus

Anjana Haridas[1], T. Makeshkumar[2] and J. Sreekumar[2]

[1] College of Agriculture, Vellayani, Kerala Agricultural University, Thiruvananthapuram 695 522, Kerala, India

[2]ICAR-Central Tuber Crops Research Instituite, Thiruvananthapuram 695 017 Kerala, India

Corresponding author: J. Sreekumar, email: Sreekumar.J@icar.gov.in

Abstract

Cassava Mosaic Disease caused by cassava mosaic virus is one of the most devastating crop diseases affecting cassava cultivation. Bioinformatics approach was applied to search cassava (*Manihot esculenta* Crantz.) miRNAs that targeted the cassava mosaic virus genome. Nucleotide sequences representing the genome of cassava mosaic virus were screened against a set of mature cassava miRNAs. Efficacy of cassava miRNAs against putative viral mRNA targets was analysed based on complementarity of miRNA-mRNA target pairing. This study revealed 14 cassava miRNA families to have putative targets in the cassava mosaic virus with nearly perfect complementarity. These miRNAs when artificially designed may have the potential to confer effective resistance against cassava mosaic disease infection in transformed cassava.

Key words: Cassava, micro RNA, cassava mosaic virus, bioinformatics

Introduction

Cassava (*Manihot esculenta* Crantz.) is a dietary staple and significant cash crop. Farmers choose it for its high productivity and capacity to produce even under adverse environmental conditions. These factors make cassava an ideal food security crop. However, cassava is susceptible to a range of biotic stresses. Cassava Mosaic Disease (CMD) is one such biotic stress. CMD caused by cassava mosaic geminiviruses results in withering of the cassava leaves and limits the root growth (Legg and Fauquet, 2004; Legg et al., 2015). Cassava Mosaic Virus belongs to the genus *Begomovirus* and family *Geminiviridae*. As on today eleven different cassava mosaic viruses reported to cause CMD across globe and in India. CMD causes losses between 20-80% of the total yield and sometimes results in complete crop failure (Malathi et al., 1985; Fregene and Puonti-Kaerlas, 2002). CMD was first noticed in India in 1942 (Abraham, 1956) and its overall incidence was higher in the two main cassava growing states of Kerala and Tamil Nadu (Mathew, 1989; Antiha et al., 2008 & 2011). The primary spread of the disease is through infected planting material and secondary spread is by an insect vector, whitefly (*Bemisia tabaci* Genn.) (Antony et al., 2006).

Geminiviruses are large family of plant viruses with circular, single stranded genomes packaged within geminate particles. Geminivirus replication relies on DNA intermediates and takes place within the nucleus via two stages: by converting the genomic ssDNA into a dsDNA intermediate and amplification of viral ssDNA through the rolling-circle replication (Gutierrez et al., 2004). The genomic ssDNA is then transported to neighbouring cells and is encapsidated to form mature viral particles.

The genome of cassava mosaic geminiviruses contains two DNA molecules – A and B each of about 2.8 kbps (Stanley, 2004) which are coding for different proteins responsible for different functions in the infection process. Both the DNA molecules are required for infectivity, vector transmission, virus spread and for the systemic infection of susceptible host plants (Fregene

and Puonti-Kaerlas, 2002). DNA A and B are involved in the replication (both DNA components and virus) and viral spread (cell-cell and long distance) respectively. DNA A encodes two overlapping virion-sense open reading frames (ORFs) AV2 and AV1, and atleast four overlapping complementary-sense ORFs AC1, AC2, AC3 and AC4. AV1 encodes the coat protein gene (CP) and is the determinant of vector transmission (Harrison et al., 2002). ORF AC1 encodes a replication-associated protein (Rep), AC2 a transcriptional activator protein (TrAP) and AC3 a replication enhancer protein (REn). ORF AC4 plays a role as a host activation protein, which serves as an important symptom determinant implicated in cell-cycle control and may also counteract the host response to Rep gene expression (Hull, 2002). The two ORFs in DNA B are non-overlapping and code for genes that play a role in intra- (BV1) and inter- (BC1) cellular movement of virions within the host plant cell (Stanley et al., 2005).

Recombination and pseudo recombination between cassava mosaic geminiviruses give rise to different strains and members of novel virus species with increased virulence causing severe disease epidemics. Nine species of CMGs have been identified between Africa and South Asia based on their genomic sequence and phylogenetic analysis. They include representatives of seven African and two south Asian species namely *African Cassava Mosaic Virus* (ACMV), *East African Cassava Mosaic Virus* (EACMV), *East African Cassava Mosaic Cameroon Virus* (EACMCV), *East African Cassava Mosaic Kenyan Virus* (EACMKV), *East African Cassava Mosaic Malawi Virus* (EACMMV), *East African Cassava Mosaic Zanzibar Virus* (EACMZV), *South African Cassava Mosaic Virus* (SACMV), all from Africa as well as *Indian Cassava Mosaic Virus* (ICMV) and *Sri Lankan Cassava Mosaic Virus* (SLCMV) in Asia (Fauquet et al., 2008; Patil and Fauquet, 2009; Legg et al., 2015).

Various approaches are currently being applied to mitigate the CMD constraint to achieve better cassava varieties. One of the strategies is the potential application of the knowledge of microRNAs (miRNAs) in gene regulation. MicroRNAs are non-coding short RNAs that negatively regulate gene expression at the post transcriptional levels by either binding to the mRNAs for degradation, or by inhibiting protein translation or by both. In plants, many miRNA targets transcription factors, which in turn regulate specific subsets of genes and hence also is designated as 'regulator of regulators' (Ambros, 2004). Plant microRNAs show nearly perfect pairing with target mRNAs. They are responsible for the silencing of the invading viral mRNA. The existence of virus-specific plant miRNAs is crucial determinant for plants to acquire virus resistance and miRNAs targets the genes of cassava mosaic virus. Bioinformatics tool reliably identified the miRNA targets in the DNA A and B of cassava mosaic virus.

Materials and Methods

A set of 153 known miRNA sequence of *M. esculenta* were downloaded from miRBase (http://www.mirbase.org/). Further 75 nucleotide sequence of DNA A and 18 nucleotide sequence of DNA B of cassava mosaic virus was retrieved from NCBI GenBank (http://www.ncbi.nlm.nih.gov/). The nucleotide sequence composed of sequences from all 9 strains of cassava mosaic virus.

Sequence and structure homologies are used for computer-based predictions of miRNAs. Computational strategies provide a valuable and efficient manner to predict miRNA genes and their targets. To identify miRNA target sites in DNA A and DNA B, miRanda target prediction algorithm (http://www.microrna.org/) was applied. Analysis was performed on Mac (OS) based computer having Intel ® core i7, 2.8 GHz processor and 4 GB RAM. The threshold sequence complementarity score was tuned at 50, free energy was adjusted at -20 kcal mol^{-1} and threshold percentage complementarity between miRNA-target duplex was selected at 60. True regulatory targets were scrutinized on the assumption that all potential miRNA targets do not have more than five mismatches. These include one or two mismatches in octameric seed region, not more than three mismatches in position 13-22 and complementarity or wobble at position 10 and 11 and not more than a single gap inserted with their corresponding miRNAs.

Results and Discussion

Bioinformatics approach is applied to identify potential cassava miRNA regulated strains of cassava mosaic virus namely EACMV, ACMV isolate West Kenyan 844 segment, EACMZV, EACMCV, EACMVKV, ICMV, SLCMV and SACMV were downloaded from NCBI GenBank and used as target transcripts. Mature cassava

miRNA from miRBase served as query in miRanda target prediction algorithm.

Potential regulatory targets, having 5 or fewer mismatches and with no gaps in full length nucleotide sequences were identified. Fourteen miRNA families, namely mes-miR159, mes-miR164, mes-miR167, mes-miR168, mes-miR171, mes-miR319, mes-miR394, mes-miR395, mes-miR397, mes-miR408, mes-miR477, mes-miR482, mes-miR1446 and mes- miR2275 were found to have the potential to target cassava mosaic virus genome with nearly perfect complementarities. Most cassava miRNAs families (85.7 %) targeted DNA A. They include mes-miR159, mes-miR164, mes-miR167, mes-miR168, mes-miR171, mes-miR319, mes-miR395, mes-miR397, mes-miR408, mes-miR477, mes-miR1446 and mes-miR2275. AC1 gene was targeted by 16 different miRNAs, followed by AV1/AC4, AC3 and AC2/AV2 which showed 12, 2 and 1 putative target, respectively. BV1 of DNA B was targeted by 7 miRNAs while BC1 was targeted by no miRNA. The miRNA families mes-miR159 and mes-miR164 had targets in both DNA A and DNA B.

East African Cassava Mosaic Zanzibar Virus genome has potential targets for 16 cassava miRNAs followed by *East African Cassava Mosaic Kenyan Virus* (13), *Indian Cassava Mosaic Virus* (11) and *Sri Lankan Cassava Mosaic Virus* (8). The remaining 5 strains are targeted by less than 5 cassava miRNAs. ICMV and SLCMV strains present in India were targeted at BV1 by mes-miR164d. The average free energy, miRanda score and percent of sequence complementarity between miRNA-target duplex are -22.01 ± 1.02 kcal/mol, 112.23 ± 13.42 and 80.31 ± 3.51 per cent respectively. The cassava miRNA - cassava mosaic virus genome interaction is summarized in Table 1 and the detailed properties of the miRNAs are given in Table 2. For experimental validation cassava miRNAs (mes-miR164 and mes-miR395) having target in cassava mosaic virus genome was randomly selected. mes-miR395 and mes-miR164 had targets in cassava mosaic virus DNA A and DNA B respectively. The expression of miRNAs was detected using a two step process. In the first, stem loop RT primers is hybridized to miRNA molecule and then reverse transcribed in a pulsed reverse transcription reaction. In the second step the RT product is amplified and quant ified using SYBR Green I assay). There was a slight increase in mes-miR395 expression in healthy leaf sample compared to cassava mosaic disease infected (Fig. 1).

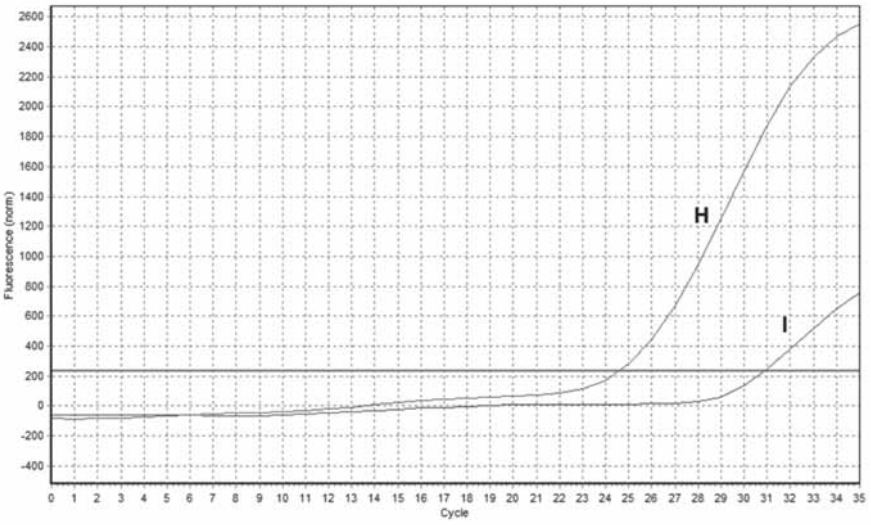

Fig. 1. mes-miR395 qPCR amplification in healthy and infected cassava cDNA samples

Table 1. Cassava miRNA targets in cassava mosaic virus

	AV1	AV2	AC1	AC2	AC3	AC4	BV1	BC1
ACMV	164d;408;1446	-	-	-	-	-	482	-
EACMV	1446;408	2275	-	-	-	-	-	-
EACMZV	168a;171g,h,i,j,k	2275	395a,b,c,d	-	-	-	159c,d;394c	-
EACMCV	408	2275	-	395e	395e	-	-	-
ICMV	319h;408	-	-	-	-	-	164d	-
SLCMV	319f,g	-	477a,b,c,d,e	-	-	-	164d	-
EACMKV	168a	2275	319a,b,c,d,e,f,g;408	-	-	-	159c,d;394a,b	-
EACMMV	-	-	-	-	397	-	-	-

Table 2. Properties of Cassava miRNAs identified by homolog search which have targets in cassava mosaic virus

mes-miR	MATURE SEQUENCE	SCAFFOLD	ML	ARM	PL	(G+C)%	(A+U)%	A/U	MFE	AMFE	MFEI	FRAME
159c	AUUGGAGUGAAGGGAGCUCUG	3581	21	3'	183	51.37	48.63	0.85	-85.6	-46.78	0.91	-
159d	AUUGGAGUGAAGGGAGCUCUG	2658	21	3'	215	47.91	52.09	0.75	-93.3	-43.40	0.91	-
164d	UGGAGAAGCAGGGCACAUGCU	6700	21	5'	84	46.43	53.57	0.96	-49.3	-58.69	1.26	+
168a	UCGCUUGGUGCAGGUCGGGAA	2477	21	5'	173	48.55	51.45	0.71	-81.2	-46.94	0.969	+
171g	UGAUUGAGCCGUGCCAAUAUC	9876	21	3'	75	44.00	56.00	0.83	-36.7	-48.93	1.11	+
171h	UGAUUGAGCCGUGCCAAUAUC	9683	21	3'	113	38.94	61.06	0.64	-46.7	-41.33	1.06	-
171i	UAUUGGCCUGGUUCACUCAGA	7520	21	3'	140	40.70	59.30	0.54	-60.8	-43.43	1.07	+
171j	UGAUUGAGCCGUGCCAAUAUC	6512	21	3'	132	44.70	55.30	0.52	-49.9	-37.8	0.85	+
171k	UGAUUGAGCCGUGCCAAUAUC	1701	21	3'	90	42.22	57.78	0.86	-42.2	-46.89	1.11	+
319f	UUGGACUGAAGGGAGCUCCUU	6446	21	3'	178	48.88	51.12	0.98	-93.5	-52.53	1.07	-
319g	UUGGACUGAAGGGAGCUCCUU	3454	21	3'	188	46.81	53.19	0.85	-84.6	-45.00	0.96	-
319h	CUUGGACUGAAGGGAGCUCCU	2877	21	3'	180	46.11	53.89	0.73	-79.0	-43.89	0.95	-
394a	UUGGCAUUCUGUCCACCUCC	9683	20	5'	151	45.70	54.30	0.64	-69.4	-45.96	1.01	+
394b	UUGGCAUUCUGUCCACCUCC	2943	20	5'	98	43.88	56.12	0.45	-43.6	-44.49	1.01	+
394c	UUGGCAUUCUGUCCACCUCCAU	6914	22	5'	93	41.94	58.06	0.86	-54.0	-58.06	1.40	-
397	UUUGAGUGCAGCGUUGAUGA	5338	21	5'	81	34.57	65.43	0.66	-27.4	-33.83	0.98	-
408	AUGCACUGCCUCUUCCCUGGC	80	21	3'	134	50.00	50.00	1.03	-65.7	-49.03	0.98	-
482	UCUUCCCUACUCCACCCAUUCC	2264	22	3'	96	45.83	54.17	0.58	-58.1	-60.52	1.32	+
1446	UUCUGAACUCUCUCCCUCAU	12262	20	5'	100	40.00	60.00	1.31	-38.7	-38.70	0.97	+

A bioinformatics approach was applied to identify endogenous *M. esculenta* miRNAs having anti-cassava mosaic disease defense by targeting DNA A and B of various strains of cassava mosaic virus. DNA A and B were targeted at several loci by various miRNAs. The degree of complementarity determines the fate of a target site. Perfect complementarity leads to endonucleolytic cleavage, while imperfect complementarity results in translational repression leading to destabilization of miRNA (Baek et al., 2008; Selbach et al., 2008).

The targeted regions in DNA A were mainly associated with replication (AC1), silencing suppressor (AC4) and coat protein (AV1). AC1 is involved in replication of viruses within the host cells. The open reading frame (ORF) encodes a replication associated proteins (Hull, 2002). Targeting AC1 gene would impact viral replication by reducing viral DNA accumulation in host. AC1 is targeted by 16 miRNAs. Among them mes-miR395 showed maximum complementarity (85.71 %). AC4 plays a role as host activation protein which serves as an important symptom determinant implicated in cell cycle control and may also counteract the host response to replication gene expression (Hull, 2002). AC4 is targeted by 12 cassava at BV1 region by mes-miR159, mes-miR164, mes-miR394 and mes-miR482. BV1 region is responsible for coding nuclear shuttle protein (NSP) that is responsible for intra cellular movement of virions within host plant cell (Stanley et al., 2005).

This suggests that these miRNA families can be manipulated as a strategy to engineer cassava plants to defend against cassava mosaic viruses. Controlling viruses following degrading their mRNAs within a plant cell is a relatively straight forward

process. The only limitation to this is the minimum level of expression of the microRNAs. This problem can be dealt with the engineering of artificial miRNAs, thus giving a solution for the virus control.

References

Abraham, A. 1956. *Farm Bulletin No. 17. Ind. Counc. Agric. Res. New Delhi*, 20 pp. Ambros, V. 2004. The functions of animal microRNAs. *Nature*, **431**: 350-355.

Anitha J., Makeshkumar, T and Edison, S. 2008. Host range of Sri Lankan cassava mosaic virus by mechanical inoculation. *J. Root crops*, **34**(1): 21-25

Anitha J., Makeshkumar, T. and Edison,S. 2011. Survey of cassava mosaic disease in Kerala. *J. root crops*., **37**: 41-47.

Anitha Jose,. Edison,S., Makeshkumar, T. 2019. Occurrence and distribution of Sri Lankan cassava mosaic viruss in the stems of cassava. *J. root crops*, **45**: 32-40

Antony, B., Lisha,V. S., Palaniswami, M. S., Sugunan,V. S., Makeshkumar, T. and Henneberry, T. J. 2006. Bemisia tabaci (Homoptera: Aleurodidae) and Indian cassava mosaic virus tansmission. *Int, J. Trop. Insect Sci*., **26**(3): 176- 182.

Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P. and Bartel, D.P. 2008. The impact of microRNAs on protein output. *Nature*, **445** (7209): 64-71.

Fauquet, C. M., Briddon, R. W., Brown, J. K., Moriones, E., Stanley, J., Zerbini, M. and Zhous, X. 2008. Geminivirus strain demarcation and nomenclature. *Arch. Virol.*, **153**(4): 783-821.

Fregene, M. and Puonti-Kaerlas, J. 2002. Cassava Biotechnology. In: Hillocks, R. J., Thresh, J. M. and Bellotti, A. C. (eds), *Cassava: Biology, Production and Utilization*. CAB International, Wallingford, Oxon, pp. 179 – 207.

Gutierrez, C., Ramirez-Parra, E., Castellano, M. M., Sanz-Burgos, A. P., Luque, A. and Missich, R. 2004. Geminivirus DNA replication and cell cycle interactions. *Vet. Microbiol.*, **98**: 111 – 119.

Harrison, B. D., Swanson, M. M., and Fargette, D. 2002. Begomovirus coat protein: serology, variations and functions. *Physiol. Mol. Plant Pathol*., **60**: 257-271.

Hull, R. 2002. Genome Organization. In: Hull, R. (ed.), *Matthew's Plant Virology* (4th Ed.). Academic Press, London, UK, pp. 171-224.

Legg, J. P. and Fauquet, C. M. 2004. Cassava mosaic geminiviruses in Africa. *Plant Mol. Biol.*, **56**(4): 585-99.

Legg, J. P., Kumar, P. L., Makeshkumar, T., Tripathi, L., Ferguson, M., Katju, E., Ntawuruhunga, P and Cuellar, W. 2015. Cassava virus diseases: biology, epidemiology and management. *Adv. Virus Res*., **91**: 85-142.

Malathi, V. G., Nair, N. G. and Shantha, P. 1985. Cassava mosaic Disease. Central Tuber Crops Research Institute Technical Bulletin Series, **5**: 18 pp.

Mathew, A. V. 1989. Studies on Indian cassava mosaic virus disease. PhD thesis, Univ. Agric. Sciences, Bangalore, 237 pp.

Patil, B. L. and Fauquet, C. M. 2009. Cassava mosaic geminiviruses: actual knowledge and perspectives. *Mol. Plant Pathol*. **10**: 685-701.

Selbach, M., Schwanhäusser, B., Thierfelder, N., Fang, Z., Khanin, R. and Rajewsky, N. 2008. Widespread changes in protein synthesis induced by microRNAs. *Nature*, **445**(7209): 58-63.

Stanley, J. 2004. Subviral DNAs associated with geminivirus disease complexes. *Vet. Microbiol.*, **98**: 121–129.

Stanley, J., Bisaro, D. M., Briddon, R. W., Brown, J. K., Fauquet, C. M., Harrison, B. D., Rybicki, E. P., and Stenger, D. C. 2005. Geminiviridae. In: Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U. and Ball, L. A. (eds), *Virus Taxonomy VIIIth Report of the ICTV*. Elsevier/Academic Press, London, pp. 301-326.