Molecular characterization and phylogenetic analysis of betasatellite molecules associated with okra yellow vein mosaic disease in Sri Lanka

Article

Accepted Version

Title Page

Title:

Molecular characterization and phylogenetic analysis of beta-satellite molecules associated with okra yellow vein mosaic disease in Sri Lanka

Authors: Tharmila C. Jeyaseelan[1,4], Emmanuel C. Jeyaseelan[1], Devika M. De Costa[2,4], Michael W.Shaw[3]

1- Department of Botany, Faculty of Science, University of Jaffna, Jaffna, JA 40000, Sri Lanka.

2- Department of Agricultural Biology, Faculty of Agriculture, University of Peradeniya, Peradeniya, KY 20400, Sri Lanka.

3- School of Agriculture, Policy and Development, University of Reading, Reading, RG6 6BZ, United Kingdom.

4- Plant Protection Board of Study, Postgraduate Institute of Agriculture, University of Peradeniya, Peradeniya, KY 20400, Sri Lanka.

Corresponding author: Tharmila C. Jeyaseelan

E-mail address: siva.tharmi3@gmail.com

stharmila@jfn.ac.lk

Abstract

In the past few years, okra growth and yield has been greatly affected in Sri Lanka by a disease called okra yellow vein mosaic (OYVM). The disease is caused by begomovirus and associated beta-satellites. The associated beta-satellite molecule determines the development and severity of the disease. Therefore, knowledge of the sequence and diversity of the beta-satellite molecules associated with OYVM disease would assist choice of virus isolates or trial locations to use when selecting resistant varieties of okra. The present study aimed to characterize the beta-satellite DNA sequences associated with OYVM disease in Sri Lanka and determine their phylogenetic relationships. Beta-satellite DNA of six virus isolates from widely separated geographical locations were sequenced and compared with already reported begomovirus beta-satellites. The beta-satellite molecules have features common to begomovirus beta-satellite DNAs: a conserved nonanucleotide TAATATTAC, a coding sequence for the protein βC1, an adenine rich region and a satellite conserved region. er Nucleotide diversity among the isolates was relatively low ($\pi = 0.027$). Recombination events were detected at a specific region in the genome of all the isolates. The isolates shared sequence identity >96 % with bhendi yellow vein beta-satellites reported from India. This is further supported by phylogenetic analysis.

Key words:

Abelmoschus esculentus, Begomovirus, Plant virus, Virus recombination

Introduction

Yellow vein mosaic (YVM) disease of Okra (*Abelmoschus esculentus*; local names Bhendi / Vendi / Bandakka) causes heavy economic losses, among pests and diseases that affect the yield and fruit quality of okra throughout the okra cultivating regions of the world (Pun et al. 2005). Leaves of infected plants are characterized by alternate green and yellow patches, vein clearing and a conspicuous yellow network of vein. In severe cases the chlorosis may extend to the inter-veinal area and may result in complete yellowing of leaves with stunting of plants (Taware et al. 2010). The disease is caused by a geminivirus called okra yellow vein mosaic virus (OYVMV) (Jose and Usha 2003) which belongs to the genus *Begomovirus* (Fauquet and Stanley 2005). In *Begomovirus*, single stranded, circular genomic DNA is encapsulated in quasi-isometric (‘‘geminate’’) virions and transmitted by whitefly (*Bemisia tabaci*) (Briddon et al. 2008).

Most OYVMV isolates are monopartite viruses with a single strand DNA (DNA A) of helper begomovirus which associated with a single strand beta-satellite molecule. The beta-satellites have a gene which codes a protein known as βC1 (Mansoor et al. 2003). It is essential for the development of typical symptoms in infected okra plants (Jose and Usha 2003), suppression of post-transcriptional gene silencing and up-regulation of viral DNA levels *in planta* (Briddon et al. 2008). Many isolates of the virus and associated satellite molecules have been characterized in various okra growing countries (Ghosh et al. 2008; Taware et al. 2010; Venkataravanappa et al. 2011). Recent studies have revealed that the OYVM disease can be caused by a bipartite begomovirus species known as Bhendi yellow vein mosaic Delhi virus [BYVDV-IN (India: Delhi: okra)]. This has both DNA-A and DNA-B molecules, but not beta-satellites, in okra plants which nonetheless have typical symptoms (Venkataravanappa et al. 2012).

In Sri Lanka cultivated okra varieties such as TV8, MI5 and MI7 are greatly affected by the disease. Cultivation of the variety TV8 (locally known as 'Paal Vendi'), a popular variety with a great market value in North and East parts of Sri Lanka, is vanishing because of widespread disease. The variety 'Haritha' was recommended as resistant by the Department of Agriculture, but in recent seasons the variety has suffered a high disease incidence, especially in Northern Sri Lanka. So far, studies on OYVM disease in Sri Lanka have mainly focusing on screening of resistant okra varieties, determination of disease incidence. More recent studies have tried to identify the causative agent, based on partial amplification of virus DNA by PCR (Senevirathna et al. 2016).

These studies have found that an okra plant showing yellow vein mosaic disease may carry mixed infections with Bhendi yellow vein mosaic virus (BYVMV) DNA-A, Okra enation leaf curl virus (OELCuV) DNA-A, beta- and alpha-satellites (Priyavathi et al. 2016). This complex mix of virus genetic materials associated with OYVM disease suggests that identification and characterization of the causative virus and associated satellite molecules may help explain why OYVM disease has become more prevalent and severe recently. In 2016, we identified the causative agent of OYVM disease of the same samples as the begomovirus OYVMV based on whole genome sequencing of the virus isolates (Tharmila et al. 2017). The present study focused on genome sequence, organisation and phylogenetic relationship of beta-satellite molecules associated with OYVM disease in Sri Lanka.

Leaf samples of asymptomatic okra plants and okra plants showing yellow vein mosaic symptoms were collected from six different widely separated locations across Sri Lanka between May and July in 2015 (Table 1). In each location, three symptomatic and three non-symptomatic leaf samples were collected from three different farmer's field or home gardens, while the okra plants were at their post flowering stage. Total DNA was extracted from both

non-symptomatic and symptomatic okra leaf samples using a DNA extraction protocol (Ghosh et al. 2009) with some modifications.

The extracted DNAs were PCR amplified with specific primers for the begomovirus genome. Presence of begomovirus in symptomatic samples was confirmed by PCR amplification of begomovirus DNA-A using a pair of degenerate primers (Deng et al. 1994). Full length beta-satellite DNA was amplified with a pair of universal primers β01 and β02 as described by Briddon et al. (2002). Six different PCR products of beta-satellite DNA, one from each location sampled, were cloned and sequenced. The PCR products were purified using a spin column PCR purification kit (NBS Biological, Huntingdon, Cambridgeshire) and cloned into pCR®2.1 vector using a TA cloning kit (Invitrogen, USA) as described in the manufacturers' protocols. The clones were sequenced by a commercial automated Sanger sequencing service (Source Bioscience, UK). The complete nucleotide sequences of the beta-satellites were deposited in GenBank database; accession numbers are given in Table 1.

Identity searches for the sequences were carried out by using the BLAST*n* program available in the NCBI. Sequence alignments were performed using MUSCLE (Edgar 2004) and pairwise identity scores were calculated with the Species Demarcation Tool (SDT) (Muhire et al. 2014). The details of sequences retrieved from GenBank database are given in Supplementary Table 1. Phylogenetic analysis was performed in MEGA 7 (Kumar et al. 2016) using the maximum likelihood algorithm with 1000 bootstrap replications. The mean pairwise number of nucleotide differences per site (π) was estimated using DnaSP v. 5.10 (Rozas et al. 2003). Putative parental viruses and recombination breakpoints were determined using Recombination Detection Program (RDP) v. 4.0 (Martin et al. 2015). Alignments were analysed using default settings for the different methods and statistical significance was defined as a *P* value less than the Bonferroni-corrected cut-off of 0.05.

PCR-mediated amplification of DNA extracts of all symptomatic samples with specific primers for begomovirus DNA-A and beta-satellite yielded products with an approximate length of 0.52 kb and 1.3 kb respectively (Data not shown). Amplifications from non-symptomatic plants were uniformly negative with all primers. This shows the frequent association of helper virus and beta-satellite with symptomatic plants.

The complete nucleotide sequences of the six beta-satellites ranged between 1318 and 1369 bp in size (Table 1). The sequences showed all the features typical of a beta-satellite (Briddon et al. 2008), including a single ORF in a complementary sense strand which encodes a 118 amino acid protein and an adenine rich region. They also contain a nonanucleotide stem-loop structure (TAATATTAC) and a satellite conserved region (SCR).

SDT based pairwise alignment shows that the Sri Lankan okra yellow vein mosaic beta-satellites (OYVMB) had between 93.4 % to 99.2 % sequence identity. The most divergent isolate OYVMB [LK:Kan:15] had 93.4 % to 94.0 % sequence identity with the rest of the Sri Lankan OYVMB. The above isolates shared > 91 % of identity with some bhendi yellow vein beta-satellites (BYVB), bhendi yellow vein mosaic beta-satellites (BYVMB) and OYVMB[IN:Aur:08] from India. This classifies the sequences as coming from the same species, according to the recently proposed beta-satellite species demarcation threshold of 91% (https://talk.ictvonline.org/files/ictv_official_taxonomy_updates_since_the_8th_report/m/plant-official/6689). Sequence identity of Sri Lankan OYVMB isolates was < 91% with sequences of okra leaf curl beta-satellites (OLCuB), bhendi yellow vein India beta-satellites (BYVIB) and croton yellow vein mosaic beta-satellite (CrYVMB [IN_Bhu_OYBHU_06]), all of which were also reported to be associated with okra yellow vein mosaic disease in India (Venkataravanappa et al. 2011~~). Beta-satellite~~

~~isolates associated with other begomovirus diseases in Sri Lanka had even lower identity (<70 %) with the Sri Lankan OYVMB isolates.~~

~~Nucleotide diversity (π) was measured in three subpopulations of beta-satellite isolates; Sri Lankan OYVMB (n=6), all the beta-satellite isolates reported from Sri Lanka (n=11) and some selected beta-satellites associated with okra (n=20). The lowest nucleotide diversity was noticed in Sri Lankan OYVMB (π=0.027) and the highest nucleotide diversity was in the total beta-satellite isolates reported from Sri Lanka (.π=0.245).~~

Briddon et al. (2003) reported two major clusters of beta-satellites corresponding to hosts in the Malvaceae and non-Malvaceae. The Sri Lankan OYVMB clustered phylogenetically with isolates from the Malvaceae and were completely separated from previously published beta-satellite sequences from Sri Lanka, as expected (data not shown). Analysis of OYVMB and selected beta-satellites associated with okra/bhendi clearly showed the close clustering of Sri Lankan OYVMB with already reported Indian BYVB and BYVMB rather than okra enation leaf curl beta-satellites (OELCuB) and OLCuB (Figure 1). The isolates OYVMB [LK:Put:15] and OYVMB [LK:Kan:15] clustered separately from the rest of the Sri Lankan OYVMB isolates.

Genetic recombination is an important process in the evolution of viruses and it is frequent in begomoviruses and inter-species recombination is the major factor behind the emergence of new begomovirus species (Lefeuvre et al. 2007). We looked for recombination events among 27 full-length genomes of different beta satellites associated with okra and other plants. A recombination event was detected in the genome of all the test OYVMB isolates. The recombination break point was detected immediately upstream of A-rich region and SCR region. The major and minor parents were inferred to be ageratum yellow vein beta-satellite

(AYVB[LK:Age:03]) and Leucas zeylanica yellow vein beta-satellite (LZYVB[LK:Leu:06]) respectively (Table 2). These are associated with common weeds, *Ageratum* spp. and *Leucas zeylanica* respectively, in Sri Lanka. The isolate OYVMB[LK:Kan:15] showed an additional recombination event a few nucleotides downstream of the previous recombination. OYVMB[IN:Aur:08] and AYVB[LK:Age:03] were inferred to be the minor and major parents, respectively. One more recombination event was detected in the analysis, where the isolates OYVMB[LK:Jaf:15] and OYVMB[LK:Kan:15] were parents of the recombinant isolate OELCuB[IN:Jal-4/okr:15].

~~In India, the beta-satellite molecules associated with okra segregated into four groups (Venkataravanappa et al. 2011): okra leaf curl beta-satellites (OLCuB); bhendi yellow vein beta-satellites (BYVB); bhendi yellow vein India beta-satellites (BYVIB); and a surprising distinct group called croton yellow vein mosaic beta-satellite (CrYVMB). In present study, the Sri Lankan beta-satellite isolates, except OYVMB[LK:Kan:15], were more closely related to BYVB (>96 %) than rest of the three groups.~~

Begomovirus may range over long distances with little differentiation in infectivity and genome sequence. As expected, Sri Lankan isolates showed great similarities with Indian isolates. The countries are close together and agricultural products are frequently exchanged, so the isolates that associated with Sri Lankan OYVM disease might have arrived from India. Sequence diversity among the Sri Lankan OYVM beta-satellites was lower than in previously published isolates from India, possibly because the country is smaller and has less geographical and environmental variation than India, or because the disease invaded recently.

Acknowledgement

The authors are grateful to Dr. George Gibbings and Dr. Ihsan Ullah for technical assistance in Molecular genetics laboratory, School of Agriculture, Policy and Development, University of

Reading, U.K. The works in Sri Lanka was supported by University Grants Commission, Government of Sri Lanka under the project: UGC/DRIC/PG/2015(ii)/JFN/01 is duly acknowledged.

References

Briddon R, Bull S, Mansoor S, Amin I, Markham G (2002) Universal primers for the PCR-Mediated amplification of DNA β. Molecular Biotechnology 20: 315–318.

Briddon R, Brown J, Moriones E, Stanley J, Zerbini M, Zhou X, Fauquet CM (2008) Recommendations for the classification and nomenclature of the DNA- β satellites of begomoviruses. Archives of Virology 153: 763–781.

Deng D, Mcgrath PF, Robinson DJ, Harrison BD (1994) Detection and differentiation of whitefly-transmitted geminiviruses in plants and vector insects by the polymerase chain reaction with degenerate primers. Annals of Applied Biology 125: 327–336.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792–1797.

Fauquet CM, Stanley J (2005) Revising the way we conceive and name viruses below the species level : A review of geminivirus taxonomy calls for new standardized isolate descriptors. Archives of Virology 150: 2151–2179.

Ghosh R, Paul S, Ghosh SK, Roy A (2009) An improved method of DNA isolation suitable for PCR-based detection of begomoviruses from jute and other mucilaginous plants. Journal of Virological Methods 159: 34–39.

208 Ghosh R, Paul S, Palit P, Das S, Das A, Acharyya S, Mir JI, Ghosh SK, Roy A (2008)
209 Molecular characterization and computational analysis of coat protein gene of an east
210 Indian isolate of Bhendi yellow vein mosaic virus and associated DNA-beta satellite.
211 Indian Journal of Virology 19: 173–182.

212 Jose J, Usha R (2003) Bhendi yellow vein mosaic disease in India is caused by association of
213 a DNA beta satellite with a begomovirus. Virology 305: 310–317.

214 Kumar S, Stecher G, Tamura K (2016) MEGA7 : Molecular evolutionary genetics analysis
215 version 7 . 0 for bigger datasets. Molecular Biology and Evolution 33: 1870–1874.

216 Lefeuvre P, Martin DP, Hoareau M, Naze F, Delatte H, Thierry M, Varsani A, Becker N,
217 Reynaud B, Lett JM (2007) Begomovirus 'melting pot' in the south-west Indian Ocean
218 islands: molecular diversity and evolution through recombination. Journal of General
219 Virology 88:3458–3468.

220 Mansoor S, Briddon RW, Zafar Y, Stanley J (2003) Geminivirus disease complexes : an
221 emerging threat. Trends in Plant Science 8: 128–134.

222 Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: Detection and analysis
223 of recombination patterns in virus genomes. Virus Evolution 1:vev003

224 Muhire B, Varsani A, Martin DP (2014) SDT: A virus classification tool based on pairwise
225 sequence alignment and identity calculation. PLoS One 9:e108277.

226 Priyavathi P, Kavitha V, Gopal P (2016) Complex nature of infection associated with yellow
227 vein mosaic disease in Bhendi (*Abelmoschus esculentus*). Current Science 111: 1511–
228 1515.

229 Pun KB, Doraiswamy S, Balasubramanian G (2005) Prediction of whitefly population and
230 Okra yellow vein mosaic virus disease incidence in okra. Indian Journal of Virology
231 16: 19-23.

232 Rozas J, Sanchez-Delbarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism
233 analyses by the coalescent and other methods. Bioinformatics 19:2496–2497.

234 Senevirathna HMSI, Wasala SK, Senanayake DMJB, Weerasekara D (2016) Characterization
235 and detection of yellow vein disease of okra (*Abelmoschus esculentus* (L.) Moench) in
236 Sri Lanka. Tropical Agricultural Research 27: 360–369.

237 Taware S, Thengane R, Harke S, Taware A (2010) Genome sequence analysis of DNA-A
238 from okra yellow vein mosaic virus isolates. International Journal of Current Research 8:
239 96–107.

240 Tharmila CJ, Jeyaseelan EC, Ihsan U, Wetten AC, De Costa DM, Shaw MW (2017) First
241 report on association of okra yellow vein mosaic virus with yellow vein mosaic
242 disease of okra (*Abelmoschus esculentus*) in Sri Lanka. Plant Disease 101:1335.

243 Venkataravanappa V, Reddy CNL, Jalali S, Reddy MK (2012) Molecular characterization of
244 distinct bipartite begomovirus infecting bhendi (*Abelmoschus esculentus* L .) in India.
245 Virus Genes 44: 522–535.

246 Venkataravanappa V, Reddy CNL, Swaranalatha P, Jalali S, Briddon RW, Reddy MK (2011)
247 Diversity and phylogeography of begomovirus- associated beta satellites of okra in India.
248 Virology Journal 8: 555.

249

250 **Table 1** Characteristic features of beta-satellites isolated from OYVMD from different locations in Sri Lanka

Samples collected districts in Sri Lanka	Beta-Satellites	Abbreviation	Accession number in NCBI GenBank	Length of the DNA (bp)	GenBank sequence showing highest sequence identity in nucleotide BLAST search	
Vavuniya	Okra yellow vein mosaic beta-satellite [Sri Lanka: Vavuniya: 2015]	OYVMB [LK:Vav:15]	KX174318	1318	Bhendi yellow vein India beta-satellite [India:Jalna:Jal:2009] (KJ462078)	96%
Jaffna	Okra yellow vein mosaic beta-satellite [Sri Lanka: Jaffna: 2015]	OYVMB [LK:Jaf:15]	KX174319	1334	Bhendi yellow vein India beta-satellite [India :Coimbatore:Co:2009] (KJ462077)	98%
Kandy	Okra yellow vein mosaic beta-satellite [Sri Lanka: Kandy: 2015]	OYVMB [LK:Kan:15]	KX174320	1369	Bhendi yellow vein mosaic beta-satellite [India:Guntur:OY112:2006] (GU111969)	94%
Matara	Okra yellow vein mosaic beta-satellite [Sri Lanka: Matara: 2015]	OYVMB [LK:Mat:15]	KX174321	1351	Bhendi yellow vein mosaic beta-satellite [India:Madurai:MKU-1:2014] (KR068483)	98%
Trincomalee	Okra yellow vein mosaic beta-satellite	OYVMVB [LK:Tri:15]	KX174322	1320	Bhendi yellow vein India beta-satellite [India:Jalna:Jal:2009] (KJ462078)	96%

	[Sri Lanka: Trincomalee: 2015]					
Puttalam	Okra yellow vein mosaic beta-satellite [Sri Lanka: Puttalam: 2015]	OYVMVB [LK:Put:15]	KX174323	1351	Bhendi yellow vein mosaic beta-satellite [India:Aurangabad:OY165:2006] (GU111977)	97%

251

252 **Table 2** Putative recombination events detected within beta-satellites associated with okra and beta-satellites reported from Sri Lanka, based on

253 full-length beta-satellite sequences

Event	Breakpoints		Recombinant	Parents		Methods[1]	*P*-Value[2]
	Begin	End		Major	Minor		
1	1076	1316	OYVMB[LK:Vav:15]	AYVB[LK:Age:03]	LZYVB[LK:Leu:06]	RGBMCS3	3.16E-06
2	1092	1332	OYVMB[LK:Jaf:15]	AYVB[LK:Age:03]	LZYVB[LK:Leu:06]	RGBMCS3	3.16E-06
3	1154	1367	OYVMB[LK:Kan:15]	AYVB[LK:Age:03]	LZYVB[LK:Leu:06]	RGBMCS3	3.16E-06
4	1109	1349	OYVMB[LK:Mat:15]	AYVB[LK:Age:03]	LZYVB[LK:Leu:06]	RGBMCS3	3.16E-06
5	1078	1318	OYVMB[LK:Tri:15]	AYVB[LK:Age:03]	LZYVB[LK:Leu:06]	RGBMCS3	3.16E-06
6	1109	1349	OYVMB[LK:Put:15]	AYVB[LK:Age:03]	LZYVB[LK:Leu:06]	RGBMCS3	3.16E-06
7	998	1127	OYVMB[LK:Kan:15]	OYVMB[IN:Aur:08]	AYVB[LK:Age:03]	RGBMCS3	5.93E-11
8	52	812	OELCuB[IN:Jal-4/okr:15]	OYVMB[LK:Jaf:15]	OYVMB[LK:Kan:15]	RGBMCS3	2.95E-15

254 [1] *R* RDP, *G* GeneConv, *B* Bootscan, *M* MaxChi, *C* Chimera, *S* SisScan, *3* 3SEQ

255 [2] The reported *P* values are for the methods by underline, and they are the lowest *P* values calculated for the region in question

256

257

Figure legends

Fig. 1 Colour coded pairwise identity matrix generated from 27 different beta-satellite DNAs, including 6 beta-satellites described in this work, 5 beta-satellites associated with different begomoviral diseases in Sri Lanka, 14 different beta-satellites associated with okra/bhendi and 2 other beta-satellites associated with chilli and cotton. See Supplementary Table 1 for details on the compared sequences. Each coloured cell represents a percentage identity score between two sequences. The coloured key indicates the correspondence between pairwise identities and the colours displayed in the matrix.

Fig. 2 Molecular Phylogenetic analysis by Maximum Likelihood method

The evolutionary history was inferred by using the Maximum Likelihood method based on the General Time Reversible model. The tree with the highest log likelihood (-10077.7604) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.9163)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 30 nucleotide sequences; Indian and Sri Lankan isolates are denoted IN and LK respectively within the parentheses following the isolate code, . All positions containing gaps and missing data were eliminated. There were a total of 806 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.