International Journal of Plant and Environment
Vol. 4, No. 2 (July, 2018), 41-49

In-Silico Investigation of Betasatellite Complexity in Papaya Leaf Curl Disease Complex

Saurabh Verma* **and Sangeeta Saxena***

Department of Biotechnology, Babasaheb Bhimrao Ambedkar University, Lucknow- 226024, Uttar Pradesh, INDIA.

Publication Info

Article history:
Received: 20.01.2018
Accepted: 09.03.2018
DOI: 10.18811/ ijpen.v4i02.5

Key words:
Begomovirus
Betasatellite
Generic resistance
Papaya Leaf Curl Disease Complex (PLCDC)
Recombination

Corresponding authors:
Mr. Saurabh Verma
Mobile: +91-9565427565
Email: saddybiot@gmail.com

Dr. Sangeeta Saxena
Mobile: +91-9450645342
Email: dr_sangeeta_saxena@yahoo.com

Abstract

Papaya leaf curl disease complex (PLCD) impose heavy losses to papaya farmers across India. It is mainly reported to be caused by *Papaya leaf curl virus* and associated begomoviruses carrying a DNA-A and betasatellite molecules. Both components are required for infliction of severe symptoms. The investigation into molecular diversity is necessary to devise intervention techniques against these begomoviruses. Based on the outcome of diversity, preferred resistance strategy against a disease complex should be generic in nature, which provides an advantage of a broad range resistance against a variety of plant virus causing same disease symptoms. For this purpose, we studied the molecular diversity through investigation of genetic complexity of the betasatellite genome using various computational methods and probed genetic complexity of betasatellite component. Multiple sequence alignment and recombination analysis were performed to understand the evolutionary relationship of betasatellite components with a DNA-A present as a helper genome in case of monopartite begomoviruses. The overall results indicate that the betasatellite molecules have evolved independently of DNA-A component and their association is a result of co-infection and due to inter- and intraspecific interaction with various forms of virus infection in plants. The information generated through this study has potential application in designing intervention strategies against PLCD, which is widespread in Indian sub-continent region that includes countries like India, Pakistan, Nepal and Bangladesh.

1. Introduction

Papaya, is grown mainly in tropical and sub-tropical regions of Indian sub-continent spanning from Pakistan in west to Bangladesh in east and adjoining north east and south east nations such as Thailand, Malaysia, Maldives and Indonesian group of islands (Mishra *et al*., 2007). Farmers in these areas grow papaya due to favorable environmental conditions, large scale consumption and high net income from total productivity. But, papaya plant is also amenable to tough climate, exposing it to abiotic conditions such as cold, drought, flooding and osmotic stress (Jeyakumar *et al*., 2007). Biotic stress upon papaya is one of the major concerns nowadays, especially disease caused by viruses have played an important destructive interference in total productivity in Indian sub-continent. Papaya is infected by several viruses such as begomoviruses which also infect crops such as chili, tomato, papaya, potato, cucurbits, etc. Recently, papaya crops in Brazil and Mexico were reported to be infected with, a sticky disease symptom named as Papaya Meleira virus, a new type of double stranded (ds)-RNA virus capable of causing complete loss in production (Sá Antunes *et al*., 2016). Papaya leaf curl disease complex has resulted in a complete wipe-out of papaya cultivation in northern regions of Indian sub-continent. It is reported to be caused by a group of single strand (ss)-DNA viruses, which are monopartite in nature and named as *Papaya leaf curl virus* (PaLCuV) and other associated viruses as well (Borah and Dasgupta, 2012; Singh-Pant *et al*., 2012). Recently, an *in-silico* study of papaya leaf curl disease genomic components has shown that this disease group is actually a complex, where each member virus has mutated and adapted to suit its host plant and geographical region (Verma and Saxena, 2017). Therefore, this disease is now called ‘Papaya leaf curl disease complex’ (PLCD) to accommodate the complexity in member composition and their genetic variability. Apart from DNA-A component of a monopartite begomovirus, the members of PLCD complex have been shown to be associated with betasatellite, which is a defective component containing partial sequence of DNA-A. This component contains 200 bp long intergenic region, regions containing an open reading frame (ORF) on complementary-sense strand named as c1 ORF (βC1), iterons, promoter elements and a Adenine (A) nucleotide rich sequence (Zhou, 2013). This c1 ORF has been shown to be a pathogenicity determinant and responsible for physical deformation and cellular damage of infected leaves (Tahir and Mansoor, 2011; Bhattacharyya *et al*., 2015). Therefore, due to this reason they are required for efficient infection in case of many monopartite begomovirus. PLCD complex has emerged as one such complex where 99% cases of symptoms and positive infection has been associated with the presence of these betasatellite molecules.

Betasatellite molecules genomic size is nearly half of DNA-A i.e. ~1.2-1.5kb. They are single stranded DNA particles and encapsidated within geminate particles (similar to other geminiviruses). Betasatellite molecules are reported to utilize Rep protein (DNA-A dependent synthesis) to initiate replication and transcription of βC1. Apart from being pathogenicity determinant, βC1 has also been shown to be helping in viral DNA accumulation, host RNAi suppression and helping in viral movement (Yang *et al*., 2011; Jia *et al*., 2016; Haxim *et al*., 2017). Because of the above reasons, the role of bipartite begomoviruses have been slowly taken over by monopartite begomoviruses populations in old world region. This phenomenon accounted for more than 250 betasatellite sequences being deposited in last 5-8 years. Such a large number of betasatellite components associated with DNA-A have given rise to several disease complexes in old world like Cotton leaf curl disease complex in cotton producing regions of Punjab area, spanning both India and Pakistan (Zubair *et al*., 2017a,b). Such complexes have given rise to many epidemics in past few years leading to large scale economic losses (Pita *et al*., 2001; da Silva *et al*., 2011; Moriones *et al*., 2017; Sattar *et al*., 2017). Therefore, this study was carried out to understand the genomic complexity of these betasatellites with respect to PLCD complex in Indian sub-continent. The genetic complexity arises due to genomic recombination and strand exchange phenomenon occurring due to interactions of two or more different types of viruses in same plant. This results in genetic exchange leading to formation of a new viral particle capable of causing much more diversely severe symptoms in same or in new hosts (Lefeuvre *et al*., 2009; Martin *et al*., 2011; Pearson *et al*., 2016).

2. Materials and Methods

2.1. Retrieval of genomic components associated with papaya leaf curl disease in Indian sub-continent from GenBank

The viral genome belonging to DNA-A component of PLCD complex were retrieved and annotated (Verma and Saxena, 2017). Betasatellite components were retrieved from National Centre for Biotechnology (NCBI) website (https://www.ncbi.nlm.nih.gov/) Nucleotide database (https://www.ncbi.nlm.nih.gov/nuccore/). “Betasatellite” term was used to retrieve nucleotide sequences of interest. Annotation of the sequences in FASTA format was compiled for further analysis (Table 1).

2.2. Phylogenetic analysis of betasatellite components associated with leaf curl disease

MUSCLE algorithm integrated into MEGA 6.0 software package was used to perform sequence alignment of annotated betasatellite components associated with PLCD complex (Edgar, 2004). The default parameters were selected to obtain alignment with highest accuracy and refined position information (Hall, 2013). Phylogenetic analysis for elucidation of evolutionary information related to betasatellite sequences was done using Maximum Likelihood (ML) algorithm available in MEGA 6.0 software package. ML method based on the Tamura-Nei model was used to infer evolutionary history.

2.3. Recombination analysis

A recombination-based investigation was conducted to get insight into inter-sequence interaction of DNA-A and betasatellite components causing leaf curl disease of papaya and associated weeds and crops. Recombination Detection Program ver.4 (RDP4) (http://web.cbio.uct.ac.za/ ~darren/rdp.html) was employed for such analysis. Various recombination detection modules are available in this package which independently analyze aligned set of sequences for recombination breakpoints, recombination boundaries and conclude with a p-value for each event, which predicts the probability of occurrence of that event (Martin *et al*., 2015). The default parameters were used for analysis of 51 betasatellite sequences aligned against 35 DNA-A sequences (Verma and Saxena, 2017). Events predicted by more than 4 tools were accepted and rest were rejected. All predicted events were manually screened and accepted after careful investigation of phylogenetic and recombination data inferred by various tools in RDP4.

3. Results and Discussion

3.1. Betasatellite components of PLCD complex and other associated crops and weeds

The PLCD complex is mainly composed of monopartite begomoviruses, which are associated with either betasatellite or alphasatellite or both. Since, this complex is found to be closely associated with tomato, chili and weeds infecting monopartite begomoviruses, therefore, a comprehensive list of betasatellites associated with DNA-A of the above begomoviruses was compiled (Table 1). The betasatellite molecules associated with weeds i.e. *Ageratum leaf curl virus,* tomato i.e. *Tomato leaf curl virus, Tomato yellow leaf curl virus and Tomato leaf curl Joydebpur virus,* chili i.e. *Chili leaf curl virus,* cotton i.e*. Cotton leaf curl Burewala virus,* papaya i.e. *Papaya leaf curl virus* were included in the list of 51 betasatellite molecules. Among all the above considered in our study, 39 were from in India, 10 from Pakistan and 2 were obtained from a Chinese region. Overall, the group total of betasatellites infecting papaya, tomato, capsicum, guar and other crop were 10, 11, 9, 5 and 12 respectively, out of 51 mentioned in the list (Table 1). The betasatellite molecules associated with crops and weeds other than papaya were considered in this study to evaluate their diversity and recombination potential. This helps in minimizing interference with resistance strategy to be employed in future.

3.2. Genetic variability of betasatellite in PLCD complex

All 51 betasatellite molecules DNA sequences were retrieved from NCBI in FASTA format and annotated as TBS1-TBS51 according to their taxonomic placements. Total genome alignment was performed with MUSCLE algorithm in MEGA6 software suite. The aligned DNA sequences of betasatellite molecules were subjected to phylogenetic analysis and a Maximum Likelihood tree was obtained (Fig. 1). The tree was clearly demarcated into two types of clusters viz. genetically related CLUSTER-1 (blue colored group) and genetically

Table 1: List of Betasatellite virus genomes used in this study causing leaf curl disease in papaya and associated crops and weeds.

Annotation	Satellite	Accession No.	Host	Geography	Length	c1 gene
TBS 1	Ageratum leaf curl betasatellite clone beta	JX512904.2	*Amaranthus hypochondriacus* L.	Lucknow, India	1362	180-596
TBS 2	Ageratum leaf curl betasatellite isolate Jaipur	KY089034.1	*Helianthus* sp.	Jaipur, India	1366	180-596
TBS 3	Ageratum leaf curl betasatellite isolate NBRI-B1	KR922821.1	*Calendula officinalis*	Lucknow, India	1363	180-596
TBS 4	Ageratum leaf curl betasatellite isolate Sikar M 2	KC589700.1	*Tagetes patula*	Lakshmangarh, Rajasthan, India	1335	180-608
TBS 5	Chilli leaf curl betasatellite -[Pakistan:Potato:2008]	FM179615.1	*Solanum tuberosum*	Punjab, Pakistan	1399	201-563
TBS 6	Chilli leaf curl betasatellite-[India:Palampur:2008]	FM877803.2	*Capsicum* sp.	Palampur, India	1376	200-562
TBS 7	Chilli leaf curl betasatellite-Panipat 4 [India:Panipat: Papaya:2008]	HM143904.1	*Carica papaya*	Panipat, Haryana, India	1369	201-557
TBS 8	Chilli leaf curl virus satellite DNA beta	EU582020.1	*Capsicum* sp.	Pataudi, India	1380	200-562
TBS 9	Chilli leaf curl virus satellite DNA beta C1 gene, isolated from Capsicum annuum in Pakistan	FN179279.1	*Capsicum annuum*	Pakistan	1387	201-563
TBS 10	Chilli leaf curl betasatellite isolate ToLCBDB-[IN:Nar:Chil:04]	JF706231.1	*Capsicum* sp.	Jodhpur, India	1380	200-562
TBS 11	Chilli leaf curl betasatellite isolate Meerut	JX193616.1	*Capsicum* sp.	Meerut, India	1390	221-670
TBS 12	Chilli leaf curl betasatellite isolate India:Punjab:TC241:2009	KJ605111.1	*Solanum lycopersicum*	Muskabad, Punjab, India	1373	201-563
TBS 13	Cotton leaf curl Burewala betasatellite	NC_013802.1	*Gossypium hirsutum*	Punjab, India	1354	198-554
TBS 14	Cotton leaf curl Burewala betasatellite, clone L2-RCA-bl-F	FN658722.1	*Gossypium hirsutum*	Punjab, India	1354	198-554
TBS 15	Papaya leaf curl virus-associated DNA beta	AY244706.1	*Carica papaya*	New Delhi, India	1372	201-557
TBS 16	Papaya leaf curl virus-associated DNA beta	NC_004706.1	*Carica papaya*	New Delhi, India	1372	201-557
TBS 17	Papaya leaf curl virus betasatellite isolate PRM	GU370715.1	*Solanum lycopersicum*	New Delhi, India	1377	200-556
TBS 18	Papaya leaf curl virus betasatellite isolate In:Var:Pum:08:1	HM101173.1	Pumpkin	Varanasi, India	1370	201-557
TBS 19	Papaya leaf curl virus betasatellite isolate PaLCuB-IYV:Del	JX050199.1	*Ipomoea purpurea*	New Delhi, India	1367	201-557
TBS 20	Papaya leaf curl virus betasatellite isolate PaLCuB-Pumpkin:IARI	JX040472.1	Pumpkin	New Delhi	1367	201-557
TBS 21	Papaya leaf curl virus betasatellite clone BG-CBE beta	KC959933.1	Black gram	Coimbatore, India	1351	188-556
TBS 22	Papaya leaf curl virus betasatellite clone GG-CBE	KC959934.1	Green gram	Coimbatore, India	1358	188-556
TBS 23	Papaya leaf curl virus beta-satellite clone BG-VBN	KC959935.1	Black gram	Pudukottai, India	1359	188-556

TBS 24	Papaya leaf curl virus beta-satellite isolate India:Pune: TC255:2010	KJ605112.1	*Solanum lycopersicum*	Pune, India	1367	200-556
TBS 25	Papaya leaf curl virus beta-satellite isolate India:Bangalore: TC281:2010	KJ605113.1	*Solanum lycopersicum*	Bangalore, India	1367	201-557
TBS 26	Papaya leaf curl betasatellite complete sequence, isolate Palampur	LN831955.1	*Valeriana jatamansi*	Palampur, India	1367	201-557
TBS 27	Papaya leaf curl betasatellite isolate NBRI	JX987089.2	*Parthenium hysterophorus* L.	Lucknow, India	1367	189-557
TBS 28	Papaya leaf curl betasatellite isolate GMT32	KT948074.1	*Cucurbita pepo*	Rawalpindi, Pakistan	1368	195-551
TBS 29	Papaya leaf curl betasatellite, clone Par-B1	LN906595.1	*Parthenium hysterophorus* L.	Lahore, Pakistan	1362	195-551
TBS 30	Papaya leaf curl betasatellite isolate India-Valsad-Cluster bean-2015	KT253636.1	*Cyamopsis tetragonoloba*	Valsad, Gujrat, India	1367	219-575
TBS 31	Papaya leaf curl betasatellite isolate India-Bhavnagar-Cluster bean-2015	KT253637.1	*Cyamopsis tetragonoloba*	Bhavnagar, Gujrat, India	1367	219-575
TBS 32	Papaya leaf curl betasatellite isolate DPB1	KX353621.1	*Solanum lycopersicum*	Sikar, Rajasthan, India	1367	189-557
TBS 33	Papaya leaf curl betasatellite isolate 2491	KY825245.1	*Capsicum* sp.	Islamabad, Pakistan	1331	221-577
TBS 34	Papaya leaf curl betasatellite isolate 2481	KY825246.1	*Capsicum* sp.	Islamabad, Pakistan	1369	221-577
TBS 35	Papaya leaf curl betasatellite isolate 49sb	KY825247.1	*Capsicum* sp.	Islamabad, Pakistan	1368	220-576
TBS 36	Tomato leaf curl virus-associated DNA beta	NC_004715.1	*Solanum lycopersicum*	Jabalpur, Madhya Pradesh, India	1424	195-221
TBS 37	Tomato leaf curl betasatellite isolate ToLCuB[IN:Bah]	KM201278.1	*Solanum lycopersicum*	India	1353	191-571
TBS 38	Tomato leaf curl betasatellite - [India:Halwadni:2007]	EU847239.1	*Solanum lycopersicum*	Haldwani, Uttarakhand, India	1370	200-556
TBS 39	Tomato leaf curl betasatellite-Naj 1 [India:New Delhi:Papaya:2008]	HM143909.1	*Carica papaya*	Najafgarh, New Delhi, India	1369	199-555
TBS 40	Tomato leaf curl betasatellite-DU [India:New Delhi:Papaya:2009]	HM143910.1	*Carica papaya*	New Delhi, India	1370	201-557
TBS 41	Tomato leaf curl betasatellite-Naj 2 [India:New Delhi:Papaya:2009]	HM143911.1	*Carica papaya*	Najafgarh, New Delhi, India	1370	201-557
TBS 42	Tomato leaf curl betasatellite-Panipat 1 [India:Panipat:Papaya:2008]	HM143901.1	*Carica papaya*	Panipat, Haryana, India	1369	201-557
TBS 43	Tomato leaf curl betasatellite-Panipat 2 [India:Panipat:Papaya: 2008]	HM143902.1	*Carica papaya*	Panipat, Haryana, India	1370	200-556
TBS 44	Tomato leaf curl betasatellite-Panipat 5 [India:Panipat:Papaya: 2008	HM143905.1	*Carica papaya*	Panipat, Haryana, India	1373	201-557
TBS 45	Tomato leaf curl betasatellite-Panipat 7 [India:Panipat:Papaya :2008	HM143907.1	*Carica papaya*	Panipat, Haryana, India	1375	201-557
TBS 46	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D17	LT009401.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557
TBS 47	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D1	LT009403.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557

TBS 48	Tomato leaf curl betasatellite complete sequence, isolate Cluster bean, clone NGS-D22	LT009406.1	*Cyamopsis tetragonoloba*	Bhakkar, Pakistan	1371	201-557
TBS 49	Tomato yellow leaf curl China virus-associated DNA beta complete genome, isolate Y87	AJ457818.1	*Solanum lycopersicum*	Yunnan, China	1344	209-565
TBS 50	Tomato yellow leaf curl China virus-associated DNA beta complete genome, isolate Y38	AJ420315.2	*Solanum lycopersicum*	Yunnan, China	1338	209-565
TBS 51	Tomato leaf curl Joydebpur beta virus	NC_010236.1	*Capsicum* sp.	Kalyani, West Bengal, India	1370	193-573

c1= betasatellite ORF

distant CLUSTER-2 (Red to magenta colored groups). Chinese taxon is illustrated separately in light blue color and outgroups in black colored labels. CLUSTER-1 was further grouped into Group1-1 to Group1-5. The CLUSTER-2 was not sub-divided into sub-clades, thus, analyzed separately.

CLUSTER 1-1 comprises of betasatellite components infecting papaya in New Delhi and Panipat regions (India) and Punjab (Pakistan). CLUSTER 1-1 was found to be genetically very close i.e. phylogenetic likelihood predicted in 80-100% bootstrap steps. TBS7, 39, 40, 41, 42, 43, 44 and 45 were

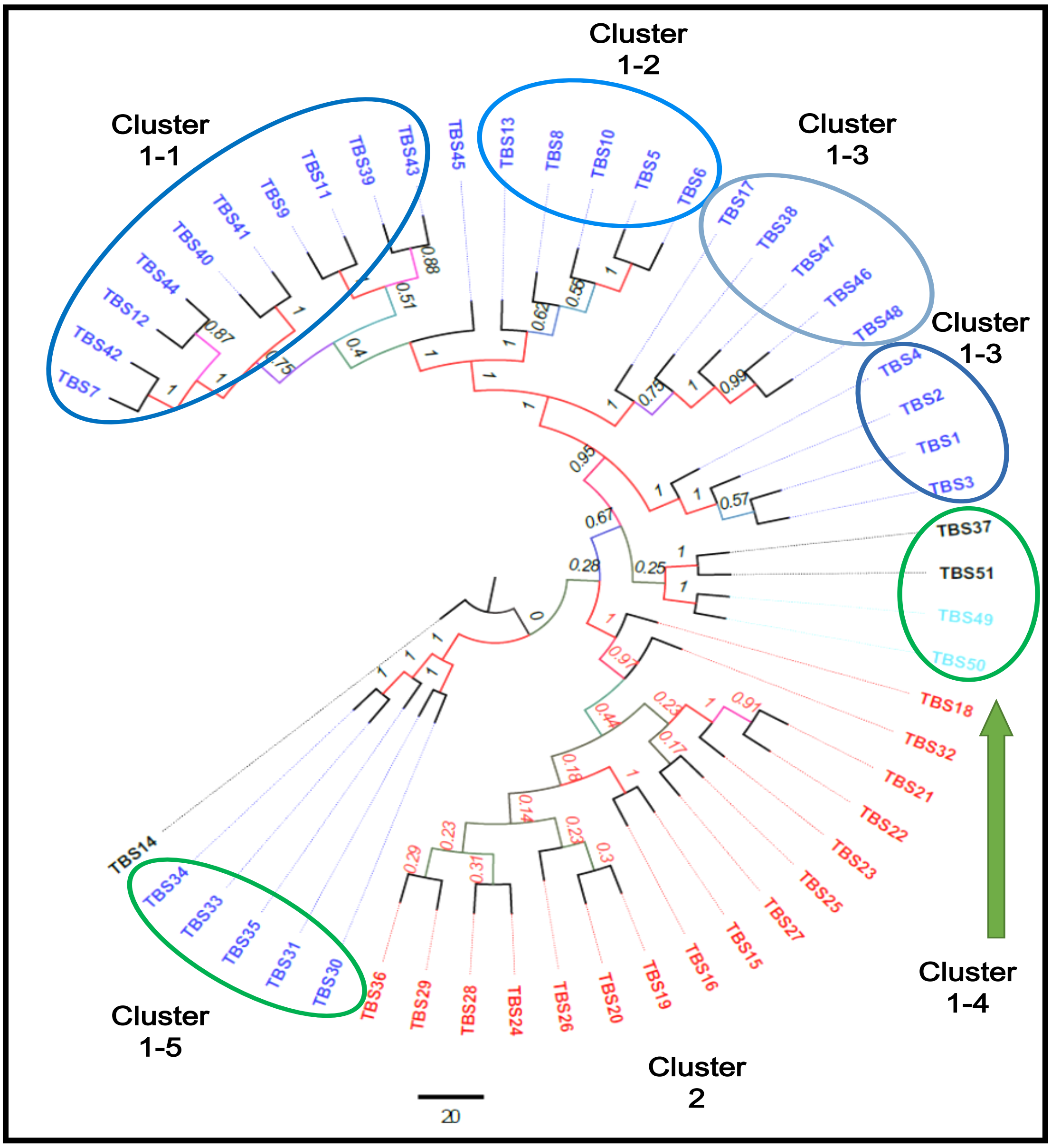

Fig. 1: Phylogenetic analysis of betasatellite component of PLCD complex and associated leaf curl disease group. The tree was divided into two clusters i.e. Cluster-1 and Cluster 2 on the basis of likelihood predicted by bootstrap step estimation. The Cluster-1 was further subdivided into five sub-clusters on the basis of host, virus and region. Cluster-2 is sparsely related, therefore, all genomic components in this group were analyzed separately. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The tree with the highest log likelihood (-13451.7843) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.5473)). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 9.4207% sites). The analysis involved 51 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 878 positions in the final dataset. Evolutionary analyses.

papaya-infecting betasatellites associated with Chili leaf curl and Tomato leaf curl disease causing DNA-A component. CLUSTER 1-2 consists of Chili leaf curl associated betasatellites infecting capsicum and tuber plants in Punjab (Pakistan), Palampur (India), Rajasthan (India) and Punjab (India) regions. CLUSTER 1-3 consists of tomato and guar plants in New Delhi, Haldwani in India and Bhakkar in Pakistan. These betasatellites were found to be associated with primarily Guar and Tomato leaf curl disease. CLUSTER 1-4 is a group of weed infecting betasatellite infecting weeds with Ageratum leaf curl disease reported mainly in Central parts of Uttar Pradesh and Rajasthan, India. CLUSTER 1-5 was found to be linked to Papaya leaf curl disease causing DNA-A components infecting guar and capsicum plants in western provinces of India and adjoining areas in Pakistan. TBS 13 is an outlier in this group belonging to betasatellite associated with Cotton leaf curl Burewala disease from Punjab (India) region.

CLUSTER 2 was observed to have complex clade pattern consisting of papaya leaf curl associated betasatellite components reported from various parts of India. The betasatellite components in this group were cladistically distant relatives as evident from low bootstrap predictions. These betasatellite components infect various crops such as papaya, tomato, cucurbits, weeds and other herbaceous plants growing in tropical regions of India and Pakistan. The papaya-infecting group consist of TBS15 and 16 while tomato was found to be infected by TBS24, 25, 32 and 36. A unique group of TBS21, 22 and 23 isolates reported from Coimbatore and Pudukottai districts (Tamil Nadu, India), infect various gram (Cereal) varieties in that region. TBS18, 19, 20 and 28 were found to be infecting cucurbits and tuber varieties. The weed infecting group comprises of TBS27 and 29 infecting *Parthenium* specifically.

Table 2: Recombination analysis of Betasatellite genomes reported in Indian sub-continent using RDPv4.95 program.

Event No.	**Recombinant**	**Major parent**	**Minor Parent**	**Frequency of recombination event**	**Detection method**						
					R	**G**	**B**	**M**	**C**	**S**	**T**
1	TBS32	TBS3	TBS18	2	+	+	+	+	+	+	+
2	TBS8	TBS10	TBS11	1	+	+	+	+	+	+	+
3	TBS4	TBS1	TBS34	1	+	+	+	+	+	+	+
4	TBS29	TBS18	TBS47	15	+	+	+	+	+	+	+
5	TBS27	TBS32	TBS15*	1	+	+	+	+	-	+	+
10	TBS27	TBS21	TBS14	13	+	+	-	+	-	+	+
11	TBS38	TBS39*	TBS6	5	+	+	-	+	+	+	+
12	TBS3	TBS2	TBS1	1	-	+	+	+	+	+	+
13	TBS35	TBS2	TBS1	4	+	+	+	+	+	+	-
15	TBS16	TBS26	TBS27*	2	+	+	+	+	+	-	-
57	TBS36	TBS27	TBS26	5	+	+	+	+	+	+	+
Total = 11				**Total frequency = 50**							

*Unknown parent detected in recombination analysis but the specified parent was the closest inferred sequence in this analysis

The analysis was performed against Betasatellite and DNA-A sequences of 51 begomoviruses using default parameters and following algorithms: R=RDP; G=GENCONV; B=BOOTSCAN; M=MAXCHI; C=CHIMAERA; S=SISCAN; T=3SEQ

From above phylogeny of betasatellite molecules, it is clear that these components have evolved in a mixed environment, interacting with more than one type viral components. Due to this reason, these defective partial viral components i.e. betasatellites, have evolved to infect a variety of host plants. This is evident from the above cladistic analysis that hosts from different families have played vital part in evolution related to different betasatellite molecules, expanding their host range and helping in their persistent presence in the region throughout year and across seasons. The reason for their omnipresent nature needs deeper investigation using recombination analysis with DNA-A components.

3.3. Recombination pattern in betasatellite component of Asian PLCD complex

A recombination analysis was carried out to investigate the genomic breakpoints in begomoviral sequences, which provide an evidence of recombination pattern prevalent between the DNA-A and betasatellite components of PLCD and other leaf curl causing complex. DNA-A dataset of PLCD complex used to investigate recombination pattern within DNA-A components was aligned with betasatellite dataset to prepare input for RDP4 software based recombination analysis. The output from various algorithm was analyzed for authenticity of analysis for breakpoint distribution and recombinants using RDP, GENCONV, MAXCHI, CHIMAERA and analysis of breakpoint boundary distribution with reference to TBS1 genome sequence (Fig. 2).

The recombination analysis of betasatellite component did not predict any recombination boundary, which includes genomic fragments from DNA-A components of PLCD. However, eleven recombination events were detected with statistical relevance and qualifying the criteria set initially (Table 2). In total, 50 recombination signals were predicted under 11 separate events. Event no. 4 was detected in 14 different recombinants namely, TBS29, 15, 16, 19, 20, 21, 22, 23, 24, 27, 28, 32, 36 and partially evident in TBS25. Similarly, event no. 10 was detected in 11 different recombinants namely, TBS 27, 15, 16, 19, 20, 24, 26, 28, 29, 36 and partially evident in TBS 25 and 32 (Table 2). Such a prominent recombination event prediction signifies that the major and minor parents

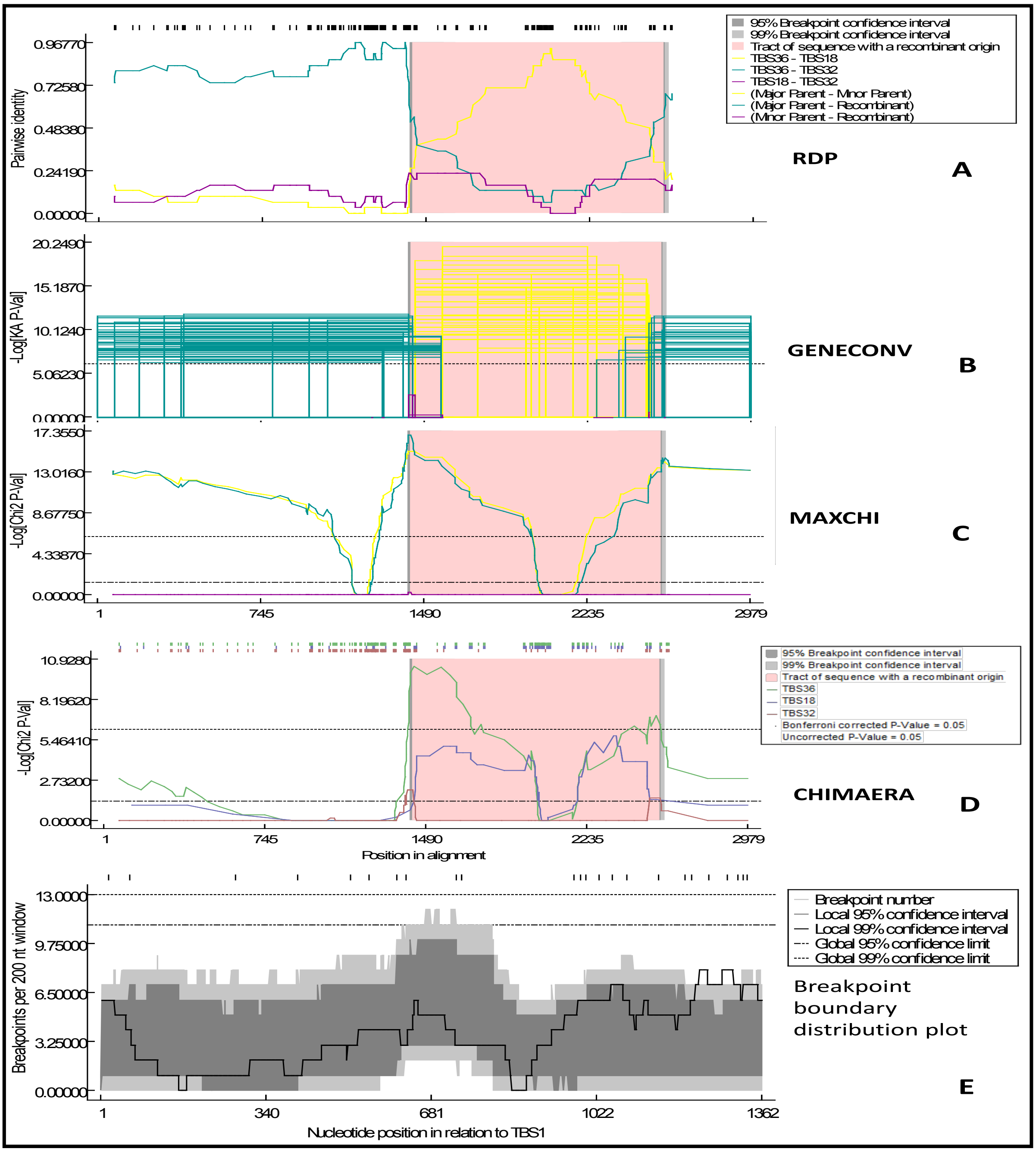

Fig. 2: Recombination analysis showing results of event number 1 using Recombination Detection Program v4.95 (RDP v4.95). The pink region indicates the occurrence of the recombination event with intersection points in the case of RDP and overlapping peaks in case of MAXCHI on the plot as the recombination starting and ending breakpoints. The grey region at the ends represents 99% and 95% confidence interval of prediction of breakpoints as mentioned in the key. A. RDP output; B. GENECONV output; C. MAXCHI output; D. CHIMAERA output; E. Recombination breakpoint distribution plot: The light and dark grey regions in the plot represent the 99% and 95% confidence limit respectively for a 200nt scanning window. The breakpoint distribution was assessed with TBS1 sequence as a reference sequence just for qualitative purpose.

Papaya leaf curl virus betasatellite isolate [In:Var:Pum:08:1] and *Tomato leaf curl betasatellite* isolate *Cluster bean*, clone *NGS-D1* respectively, were responsible for evolution of 15 putative recombinant betasatellite components reported to be associated with *Papaya leaf curl virus* DNA-A isolated from Indian tropical regions. Another pair of major and minor parents i.e. *Papaya leaf curl virus betasatellite clone BG-CBE and Cotton leaf curl Burewala betasatellite*, clone *L2-RCA-b1-F* respectively, have significant contribution in evolution of betasatellite components of above mentioned components of PLCD complex. Apart from PLCD complex, the above-mentioned parental pair has contributed in evolution of tomato leaf curl disease associated betasatellite components. Other predicted recombination events i.e. event no. 11 and 57 also gave rise to recombinants associated with tomato leaf curl disease complex (TLCD) (Moriones *et al*., 2017).

Recombination events have been reported in many cases where more than one type of betasatellite was isolated and were found to cause mixed type of symptoms in sunn hemp (Kumar *et al*., 2010) and tomato (Kumar *et al*., 2013). Another case was reported where a bipartite begomovirus i.e. *Tomato leaf curl Gujarat virus* was found to be associated with betasatellite rather than DNA-B component in *Xanthium strumarium* (a weed), which acts as a reservoir for interspecific mixing of DNA-A and associated genomic components (Mubin *et al*., 2012). These natural recombinants are proof of natural selection of factors that increase fitness of newer recombinant progenies as in betasatellite genomes. Experimental investigations have provided valuable insights into the selection of recombinant viruses which exhibit recovery of wild-type portions with enhanced fitness when compared with synthetic chimaeras, in case of mixed infections (van der Walt *et al*., 2009). Recent studies have indicated that mutations have played a pivotal role in evolution of viruses, but recombination has important role in enhancing synergistic interactions in case of mixed infections (Pita *et al*., 2001; Lima *et al*., 2017). Such synergistic adaptations often are transferred in more than one progeny, which further propagates into several other recombinant betasatellite genomes, thus, giving rise to recombination 'hot-spots' (Lefeuvre *et al*., 2007; Paul *et al*., 2016).

4. Conclusion

In this study, a significant finding that the betasatellite components associated with leaf curl disease complex of papaya and tomato have actually evolved through recombination phenomenon sheds light on the much more complex interaction among its members. This recombination has enabled these betasatellite components to evolve and infect different crops in varied tropical regions of Indian sub-continent. The broad host spectrum of PLCD and TLCD could be attributed to this phenomenon, therefore, the resistance strategies against these diseases have failed in past and are potentially prone to failure in near future. This study also explains the occurrence of interspecies infections i.e. association of Papaya betasatellite with TLCD and Tomato betasatellite with PLCD. The two disease complex make up for the large proportion of leaf curl disease in Indian sub-continent, caused by a monopartite ss-DNA viruses. Therefore, betasatellite recombination hotspots should be identified and future siRNA based strategies should be based upon these regions along-with the DNA-A component associated with respective monopartite begomoviruses.

Acknowledgment

SS and SV acknowledge Babasaheb Bhimrao Ambedkar University, Lucknow for providing the necessary infrastructure to carry out this research. SV acknowledges Babasaheb Bhimrao Ambedkar University, Lucknow for fellowship.

References

Sá Antunes, T.F. Amaral, R.J., Ventura, J.A., Godinho, M.T., Amaral, J.G., Souza, F.O., Zerbini, P.A., Zerbini, F.M. and Fernandes, P.M. 2016. The dsRNA virus papaya meleira virus and an ssRNA virus are associated with papaya sticky disease. *PLoS ONE* **11**:e0155240.

Bhattacharyya, D., Gnanasekaran, P., Kumar, R.K., Kushwaha, N.K., Sharma, V.K., Yusuf, M.A., Chakraborty, S. 2015. A geminivirus betasatellite damages the structural and functional integrity of chloroplasts leading to symptom formation and inhibition of photosynthesis. *Journal of Experimental Botany* **66**:279-281.

Borah, B.K. and Dasgupta, I. 2012. Begomovirus research in India: A critical appraisal and the way ahead. *Journal of Biosciences* **37**:791-806.

da Silva, S.J., Castillo-Urquiza, G.P., Hora Júnior, B.T., Assunção, I.P., Lima, G.S., Pio-Ribeiro, G., Mizubuti, E.S. and Zerbini, F.M. 2011. High genetic variability and recombination in a begomovirus population infecting the ubiquitous weed Cleome affinis in northeastern Brazil. *Archives of Virology* **156**:279-281.

Edgar, R.C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**:279-281.

Hall, B.G. 2013. Building phylogenetic trees from molecular data with MEGA. *Molecular Biology and Evolution* **30**:279-281.

Haxim, Y., Ismayil, A., Jia, Q., Wang, Y., Zheng, X., Chen, T., Qian, L., Liu, N., Wang, Y., Han, S., Cheng, J., Qi, Y., Hong, Y. and Liu, Y. 2017. Autophagy functions as an antiviral mechanism against geminiviruses in plants. *eLife* **6**:279-281.

Jeyakumar, P., Kavino, M., Kumar, N. and Soorianathasundaram, K. 2007. Physiological performance of papaya cultivars under abiotic stress conditions. *Acta Horticulturae* 209-216.

Jia, Q., Liu, N., Xie, K., Dai, Y., Han, S., Zhao, X., Qian, L., Wang, Y., Zhao, J., Gorovits, R., Xie, D., Hong, Y., Liu, Y. 2016. CLCuMuB βC1 Subverts Ubiquitination by Interacting with NbSKP1s to Enhance Geminivirus Infection in Nicotiana benthamiana. *PLoS Pathogens* **12**:1-30.

Kumar, A., J. Kumar, Z. A. Khan, N. Yadav, V. Sinha, D. Bhatnagar, and J. A. Khan. 2010. Study of Betasatellite Molecule from Leaf Curl Disease of Sunn Hemp (Crotalaria juncea) in India. *Virus Genes* **41**:432-40.

Kumar, J., Singh, S. P., Kumar, A., Khan, J. A., and Tuli, R. 2013. Detection and characterization of a new betasatellite: Variation in disease symptoms of tomato leaf curl Pakistan virus-India due to associated betasatellite. *Archives of Virology* **158**:257-261.

Lefeuvre, P., Lett, J.M., Varsani, A. and Martin, D.P. 2009. Widely Conserved Recombination Patterns among Single-Stranded DNA Viruses. *Journal of Virology* **83**:279-281.

Lefeuvre, P., Martin, D.P., Hoareau, M., Naze, F., Delatte, H., Thierry, M., Varsani, A., Becker, N., Reynaud, B. and Lett, J.M. 2007. Begomovirus "melting pot" in the south-west Indian Ocean islands: Molecular diversity and evolution through recombination. *Journal of General Virology* **88**:3458-3468.

Lima, A.T.M., Silva, J.C.F., Silva, F.N., Castillo-Urquiza, G.P., Silva, F.F., Seah, Y.M., Mizubuti, E.S.G., Duffy, S. and Zerbini, F.M. 2017. The diversification of begomovirus populations is predominantly driven by mutational dynamics. *Virus Evolution* 2/**3**:1.

Martin, D.P., Lefeuvre, P., Varsani, A., Hoareau, M., Semegni, J.-Y., Dijoux, B., Vincent, C., Reynaud, B. and Lett, J.-M. 2011. Complex recombination patterns arising during geminivirus coinfections preserve and demarcate biologically important intra-genome interaction networks. *PLoS Pathogens* **7**:279-281.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A. and Muhire, B. 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* **1**:1-5.

Mishra, M., Chandra, R. and Saxena, S. 2007. Papaya. In: Kole, C. (Ed) *Fruits and Nuts*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 343-351.

Moriones, E., Praveen, S. and Chakraborty, S. 2017. Tomato Leaf Curl New Delhi Virus: An emerging virus complex threatening vegetable and fiber crops. *Viruses* **9**:e264.

Mubin, M., Akhtar, S., Amin, I., Briddon, R. W., and Mansoor, S. 2012. Xanthium strumarium: A weed host of components of begomovirus-betasatellite complexes affecting crops. *Virus Genes* **44**:112-119.

Paul, P., Nag, D. and Chakraborty, S. 2016. Recombination hotspots: Models and tools for detection. *DNA Repair* (Amst). **40**:47-56.

Pearson, V.M., Caudle, S.B. and Rokyta, D.R. 2016. Viral recombination blurs taxonomic lines: examination of single-stranded DNA viruses in a wastewater treatment plant. *Peer Journal* **4**:279-281.

Pita, J.S., Fondong, V.N., Sangare, A., Ogwal, S., Fauquet, C.M. 2001. Recombination, pseudorecombination and synergism of geminiviruses are determinant keys to the epidemic of severe cassava mosaic disease in Uganda. *Journal of General Virology* **82**:655-665.

Sattar, M.N., Iqbal, Z., Tahir, M.N. and Ullah, S. 2017. The prediction of a new CLCuD epidemic in the old world. *Frontiers in Microbiology* **8**:279-281.

Singh-Pant, P., Pant, P., Mazumdar-Leighton, S. 2012. Spatial and temporal diversity of begomoviral complexes in papayas with leaf curl disease. *Archives of Virology* **157**:1217-1232.

Tahir, M.N. and Mansoor, S. 2011. C1 of chili leaf curl betasatellite is a pathogenicity determinant. *Virology Journal* **8**:279-281.

van der Walt, E., Rybicki, E.P., Varsani, A., Polston, J.E., Billharz, R., Donaldson, L., Monjane, A.L., Martin, D.P. 2009. Rapid host adaptation by extensive recombination. *Journal of General Virology* **90**:734:746.

Verma, S. and Saxena, S. 2017. Dissecting Papaya Leaf Curl Disease (PLCD) Complex and Assessing its Potential for siRNA Based Targeting. *International Journal of Plant and Environment* **3**:279-281.

Yang, X., Xie, Y., Raja, P., Yang, X., Xie, Y., Raja, P., Li, S., Wolf, J.N., Shen, Q., Bisaro, D.M. and Zhou, X. 2011. Suppression of methylation-mediated transcriptional gene silencing by βc1-sahh protein interaction during geminivirus-betasatellite infection. *PLoS Pathogens* **7**:279-281.

Zhou, X. 2013. Advances in understanding Begomovirus Satellites. *Annual Reviews of Phytopathology* **51**:357-381.

Zubair, M., Zaidi, S.S., Shakir, S., Amin, I. and mansoor, S. 2017a. An Insight into Cotton Leaf Curl Multan Betasatellite, the Most Important Component of Cotton Leaf Curl Disease Complex. *Viruses* **9**:279-281.

Zubair, M., Zaidi, S.S., Shakir, S., Farooq, M., Amin, I., Scheffler, J.A., Scheffler, B.E. and Mansoor, S. 2017b. Multiple begomoviruses found associated with cotton leaf curl disease in Pakistan in early 1990 are back in cultivated cotton. *Scientific Reports* **7**:279-281.